GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 04:31:04 ; Search time 11423 Seconds
(without alignments)
11553.842 Million cell updates/sec

Title: US-09-430-412A-1
Perfect score: 3045
Sequence: 1 atcatcaataatatccgtta.....gtggaggatggggcttctcg 3045

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	2731	89.7	139939	9	AC008965	AC008965 Homo sapi
	2	2731	89.7	180483	9	AC122707	AC122707 Homo sapi
c	3	1959.2	64.3	3000	6	AX194837	AX194837 Sequence
	4	979.4	32.2	3001	6	AX194838	AX194838 Sequence
	5	747.8	24.6	1204	9	HS5HT1A	Z11168 H.sapiens 5
	6	522.4	17.2	2512	6	AX194847	AX194847 Sequence
	7	438	14.4	699	9	HS5HT1AR5	X53826 H.sapiens D
	8	163.6	5.4	3635	10	AF087675	AF087675 Rattus no
	9	162.6	5.3	203941	2	AC114573	AC114573 Mus muscu
	10	159.6	5.2	259367	2	AC130951	AC130951 Rattus no
	11	159.6	5.2	260103	2	AC094978	AC094978 Rattus no
	12	144	4.7	5570	10	AY029704	AY029704 Mus muscu
	13	139.6	4.6	2725	10	AF217200	AF217200 Rattus no
	14	134.2	4.4	1256	10	MMU33820	U33820 Mus musculu
c	15	125.6	4.1	7218	6	I66494	I66494 Sequence 14
c	16	100	3.3	158699	9	AP006183	AP006183 Homo sapi
c	17	95	3.1	133877	2	AC120883	AC120883 Homo sapi
	18	94	3.1	129240	9	AC084128	AC084128 Homo sapi
	19	94	3.1	170928	9	AC034148	AC034148 Homo sapi
	20	92	3.0	160732	9	AC018647	AC018647 Homo sapi
c	21	91.8	3.0	169122	9	AC106803	AC106803 Homo sapi
c	22	91.6	3.0	67970	3	PFMAL1P3	AL031746 Plasmodiu
c	23	91.2	3.0	148075	9	AC110014	AC110014 Homo sapi
	24	91.2	3.0	250029	3	AE014816	AE014816 Plasmodiu
c	25	90.2	3.0	76568	3	MBREV	AF538053 Monosiga
c	26	90	3.0	139409	9	AC069181	AC069181 Homo sapi
	27	89.2	2.9	8056	6	AX599046	AX599046 Sequence
c	28	89	2.9	178207	9	AC140172	AC140172 Homo sapi
c	29	89	2.9	258658	3	AE014832	AE014832 Plasmodiu
	30	88.8	2.9	133877	2	AC120883	AC120883 Homo sapi
	31	88.6	2.9	108902	2	AC011430	AC011430 Homo sapi
c	32	88.6	2.9	108908	3	PFMAL3P8	AL034560 Plasmodiu
	33	88.4	2.9	16343	3	AMFGENOM	L06178 Apis mellif
	34	88.2	2.9	175559	9	AC112498	AC112498 Homo sapi
c	35	87.8	2.9	104992	2	AC005504	AC005504 Plasmodiu
c	36	87.8	2.9	169546	2	AC004157	AC004157 Plasmodiu
c	37	87.8	2.9	204662	9	CNS01RHQ	AL162191 Human chr
	38	87.8	2.9	250421	3	AE014849	AE014849 Plasmodiu
	39	87.6	2.9	2687	3	PFAKPA	M19881 P.falciparu
	40	87.4	2.9	159618	9	AC096750	AC096750 Homo sapi
	41	87.2	2.9	143331	9	AC091214	AC091214 Homo sapi
c	42	87.2	2.9	271546	3	AE014843	AE014843 Plasmodiu
	43	86.8	2.9	8056	6	AX598900	AX598900 Sequence
c	44	86.8	2.9	110000	2	PFMAL8P1_06	Continuation (7 of
	45	86.8	2.9	135121	9	AC069525	AC069525 Homo sapi

ALIGNMENTS

RESULT 1
AC008965/c
LOCUS AC008965 139939 bp DNA linear PRI 31-JUL-2001
DEFINITION Homo sapiens chromosome 5 clone CTD-2364L17, complete sequence.
ACCESSION AC008965
VERSION AC008965.6 GI:15042790
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 139939)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished

```
REFERENCE   2  (bases 1 to 139939)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 139939)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JUL-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 31, 2001 this sequence version replaced gi:9256099.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.5% of Sequence;
            Estimated Total Number of Errors is 0.5.
            STS Content:
            SHGC-105347 G58762
            SHGC-105921 G58845.
FEATURES             Location/Qualifiers
     source          1..139939
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2364L17"
ORIGIN

  Query Match             89.7%;  Score 2731;  DB 9;  Length 139939;
  Best Local Similarity   97.4%;  Pred. No. 0;
  Matches 2984;  Conservative    5;  Mismatches   45;  Indels   29;  Gaps   20;

Qy          1 ATCATCAATAATATCCGTTATAAAGCTTGCTTTTCTTTAGGTTAACTTTAGAGGCCTTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      95355 ATCATCAATAATATCCGTTATAAAGCTTGCTTTTCTTTAGGTTAACTTTAGAGGCCTTGA 95296

Qy         61 AGAATAAGAGCTCATCTCTTTA--CAGGAGCTTTGGTTTGCAGCATTACTTAAGAAATA 118
              ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db      95295 AGAATAAGAGCTCATCTCTTTATGCAGGAGCTTTGGTTTGCAGCATTACTTAAGAAATA 95236

Qy        119 TTTGGTATTCTGTATCTTTAAGAGTTAAACATAGAAGAATTGGCTAAGTGAAAATGAATG 178
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      95235 TTTGGTATTCTGTATCTTTAAGAGTTAAACATAGAAGAATTGGCTAAGTGAAAATGAA-G 95177

Qy        179 AAACGCAATATCATTCTGCATATATCATTTATTATATATCACAGTATTATTAGTTTT-AA 237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db      95176 AAACGCAATATCATTCTGCATATATCATTTATTATATATCACAGTATTATTAGTTTTAAA 95117

Qy        238 AAGTTAAACATAAATATCTATTATGYCATTGSACGAYTAGGYSAACCTARTCRGTGCTGC 297
              ||||||||||||||||||||||| |:|| | :  ||:|| |  | |  ||:||:|||||||
Db      95116 AAGTTAAACATAAATATCTATTA-GTCACTTGCAGACTATGGAACCATAGTCAGTGCTGC 95058

Qy        298 GAATACTTTCGATACTTCTGTTTCCCTCCTAGTATTCATAAGTGTGCCTTTGAAAACGTT 357
                |    ||  |||||||||||||||||||||||||||||||||||||||| ||||| ||
Db      95057 ATA---CTTTCATACTTCTGTTTCCCTCCTAGTATTCATAAGTGTGCCTTTAAAAAC-TT 95002

Qy        358 TTAAATTGTAAGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACTT 417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      95001 TTAAATTGTAAGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACTT 94942

Qy        418 GAATTACTTTGGATTTTGAAAAACTTTGATAAATTCTACATCATAGCATATTGAAGCAAG 477
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||| ||
Db      94941 GAATTACTTTGGATTTTTAAAAACTTTGATAAATTCTACATCATAGCATATTGAAGC-AG 94883

Qy        478 AATAACAAATGCTATACCTCAGGAATATTAATTCCAGATTTTACAGCATTTTAACTTTCT 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94882 AATAACAAATGCTATACCTCAGGAATATTAATTCCAGATTTTACAGCATTTTAACTTTCT 94823

Qy        538 TGATGAGAAAAAAATAAATTTGTCAGTTATTAAACTATTT-GGATCCAACAGATGAAAGCA 596
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db      94822 TGATGAGAAAAAAATAAATTTGTCAGTTATTAAACTATTTGGGATCCAACAGATGAAAGCA 94763

Qy        597 GAATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTT 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94762 GAATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTT 94703

Qy        657 GACACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTC 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94702 GACACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTC 94643

Qy        717 CTTTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTC 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94642 CTTTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTC 94583

Qy        777 AGCCTTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTTAACTGTAA 836
              || |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db      94582 AG-CTTTTAAATATTTTCCCATTATAATTTTTG-GACCTCTAACTCTATTTTAACTGTAA 94525

Qy        837 ATATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTT 896
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94524 ATACAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTT 94465

Qy        897 -AAAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTGC 955
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      94464 AAAAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTTC 94405

Qy        956 ATGCC----ATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTG-AATGATTGAGT 1010
              |   |    |       || |  ||| |     ||||||||||||||| |||||||||||
Db      94404 AATGCCATGAATCATCACAAATGGCAGGCTCATTGTGGTGGCATGCTGAAATGATTGAGT 94345

Qy       1011 GGGACTGTGCC----AGCTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACA 1066
              ||||      |      |||||||||||||||||||||||||||||||||||||||||||
Db      94344 GGGAACTGTGCCCAGCACTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACA 94285

Qy       1067 CACTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTG 1126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94284 CACTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTG 94225

Qy       1127 CTCTTTTGTTTGGGCTTGGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACAC 1186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94224 CTCTTTTGTTTGGGCTTGGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACAC 94165

Qy       1187 AATATTAAGATTATTGCAATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTT 1246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94164 AATATTAAGATTATTGCAATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTT 94105

Qy       1247 TTGAGATTGCATTCCTTTCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTGTG 1306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94104 TTGAGATTGCATTCCTTTCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTGTG 94045

Qy       1307 TGGTATGTTTACTGTAGTTGCTTAGAAGT-CCATTCTTTACCAATGCTCAAATGTGATTA 1365
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db      94044 TGGTATGTTTACTGTAGTTGCTTAGAAGTCCCATTCTTTACCAATGCTCAAATGTGATTA 93985

Qy       1366 AATTTG-TTTTCTTGTTAAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTTA-TT 1423
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db      93984 AATTTGTTTTTCTTGTTAAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTTATTT 93925

Qy       1424 TCAGTGATTTAACATTTCCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTTTCT 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93924 TCAGTGATTTAACATTTCCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTTTCT 93865

Qy       1484 CCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACCAAC 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93864 CCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACCAAC 93805

Qy       1544 ACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTGAAT 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93804 ACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTGAAT 93745

Qy       1604 GACAGATACTTCAGGCTTTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAAGGT 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93744 GACAGATACTTCAGGCTTTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAAGGT 93685

Qy       1664 TCAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAACTC 1723
```

```
Db      93684 TCAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAACTC 93625

Qy       1724 AATACTACCTTGTCTTTTAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCA 1783

Db      93624 AATACTACCTTGTCTTTTAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCA 93565

Qy       1784 ATACTTGCTTCATTTCTGGCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAA 1843

Db      93564 ATACTTGCTTCATTTCT-GCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAA 93506

Qy       1844 GGTGGCGAACATAAAACCTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGA 1903

Db      93505 GGTGGCG-ACATAAAACCTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGA 93447

Qy       1904 GATTAAGAGAGGCTAGCCGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTT 1963

Db      93446 GATTAAGAGAGGCTAGCCGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTT 93387

Qy       1964 GCAGGCTCTGGTAAGAAGTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACA 2023

Db      93386 GCAGGCTCTGGTAAGAAGTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACA 93327

Qy       2024 TATGCAAAATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGT 2083

Db      93326 TATGCAAAATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGT 93267

Qy       2084 CACTGAAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCA 2143

Db      93266 CACTGAAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCA 93207

Qy       2144 CACTGAGTTTTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACAC 2203

Db      93206 CACTGAGTTTTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACAC 93147

Qy       2204 GTCCTTTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAAGTCACAGGCAAT 2263

Db      93146 GTCCTTTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAAGTCACAGGCAAT 93087

Qy       2264 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 2323

Db      93086 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 93027

Qy       2324 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 2383

Db      93026 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 92967

Qy       2384 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 2443

Db      92966 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 92907

Qy       2444 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 2503

Db      92906 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 92847

Qy       2504 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 2563

Db      92846 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 92787

Qy       2564 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 2623

Db      92786 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 92727

Qy       2624 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGA 2682

Db      92726 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGA 92667

Qy       2683 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 2742

Db      92666 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 92607

Qy       2743 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 2802

Db      92606 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 92547

Qy       2803 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 2862

Db      92546 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 92487

Qy       2863 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 2922

Db      92486 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 92427

Qy       2923 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 2982

Db      92426 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 92367

Qy       2983 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 3042

Db      92366 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 92307

Qy       3043 TCG 3045

Db      92306 TCG 92304
```

RESULT 2
AC122707

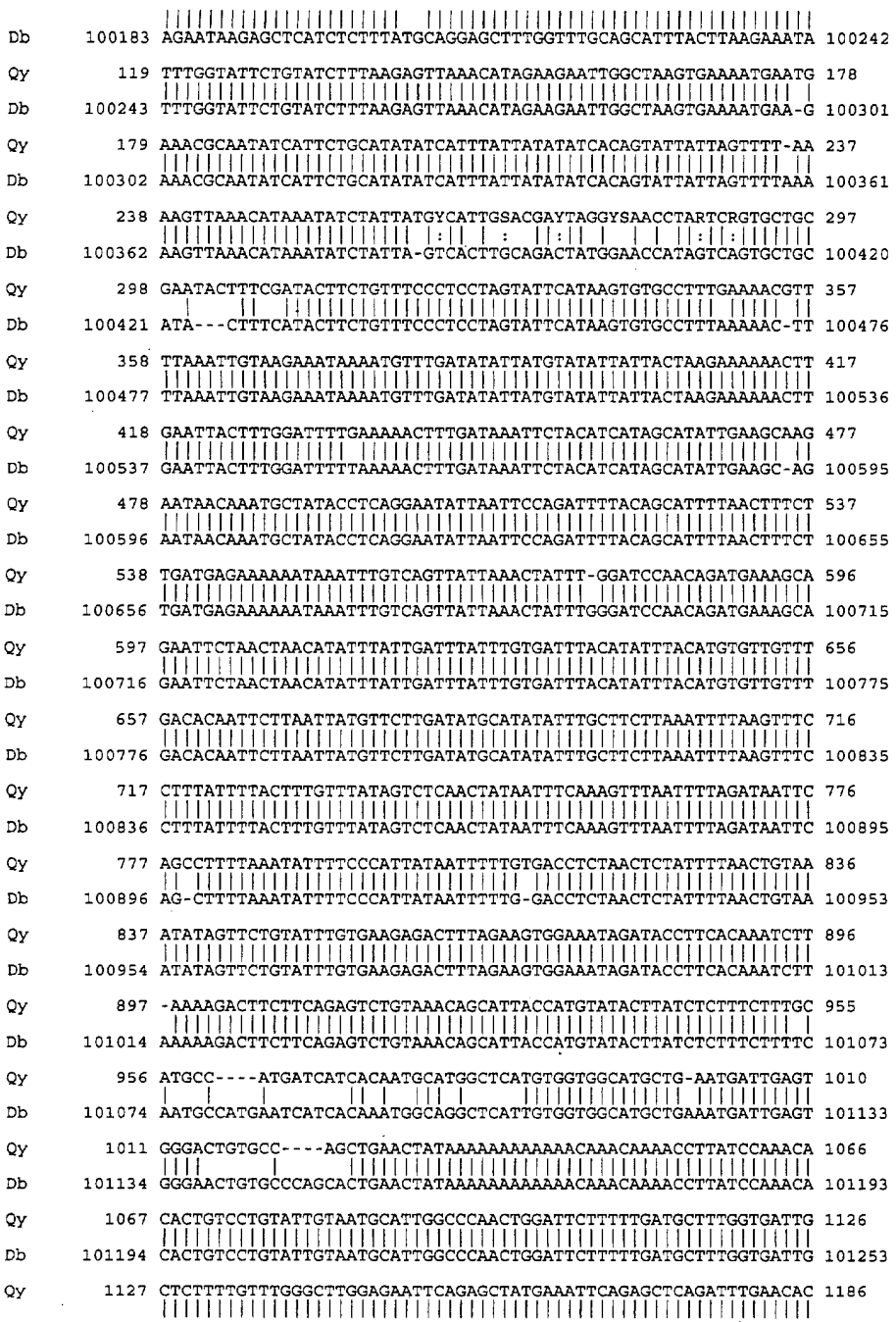
LOCUS	AC122707 180483 bp DNA linear PRI 15-JUN-2002
DEFINITION	Homo sapiens chromosome 5 clone RP11-158J3, complete sequence.
ACCESSION	AC122707
VERSION	AC122707.2 GI:21427685
KEYWORDS	HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 180483)
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 180483)
AUTHORS	DOE Joint Genome Institute.
TITLE	Direct Submission
JOURNAL	Submitted (25-MAY-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE	3 (bases 1 to 180483)
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE	Direct Submission
JOURNAL	Submitted (15-JUN-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	On Jun 15, 2002 this sequence version replaced gi:21206270. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov Finishing Completed at Stanford Human Genome Center www-shgc.stanford.edu Quality: Phrap Quality >=40 99.9% of Sequence; Estimated Total Number of Errors is 0.3.

FEATURES	Location/Qualifiers
source	1. .180483 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="5" /clone="RP11-158J3"

ORIGIN

Query Match 89.7%; Score 2731; DB 9; Length 180483;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 2984; Conservative 5; Mismatches 45; Indels 29; Gaps 20;

```
Qy          1 ATCATCAATAATATCCGTTATAAAGCTTGCTTTTCTTTAGGTTAACTTTAGAGGCCTTGA 60

Db     100123 ATCATCAATAATATCCGTTATAAAGCTTGCTTTTCTTTAGGTTAACTTTAGAGGCCTTGA 100182

Qy         61 AGAATAAGAGCTCATCTCTTTA--CAGGAGCTTTGGTTTGCAGCATTTACTTAAGAAATA 118
```


Db 100183 AGAATAAGAGCTCATCTCTTTATGCAGGAGCTTTGGTTTGCAGCATTTACTTAAGAAATA 100242
Qy 119 TTTGGTATTCTGTATCTTTAAGAGTTAAACATAGAAGAATTGGCTAAGTGAAAATGAATG 178
Db 100243 TTTGGTATTCTGTATCTTTAAGAGTTAAACATAGAAGAATTGGCTAAGTGAAAATGAA-G 100301
Qy 179 AAACGCAATATCATTCTGCATATATCATTTATTATATATCACAGTATTATTAGTTTT-AA 237
Db 100302 AAACGCAATATCATTCTGCATATATCATTTATTATATATCACAGTATTATTAGTTTTAAA 100361
Qy 238 AAGTTAAACATAAATATCTATTATGYCATTGSACGAYTAGGYSAACCTARTCRGTGCTGC 297
Db 100362 AAGTTAAACATAAATATCTATTA-GTCACTTGCAGACTATGGAACCATAGTCAGTGCTGC 100420
Qy 298 GAATACTTTCGATACTTCTGTTTCCCTCCTAGTATTCATAAGTGTGCCTTTGAAAACGTT 357
Db 100421 ATA---CTTTCATACTTCTGTTTCCCTCCTAGTATTCATAAGTGTGCCTTTAAAAAC-TT 100476
Qy 358 TTAAATTGTAAGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACTT 417
Db 100477 TTAAATTGTAAGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACTT 100536
Qy 418 GAATTACTTTGGATTTTGAAAAACTTTGATAAATTCTACATCATAGCATATTGAAGCAAG 477
Db 100537 GAATTACTTTGGATTTTTAAAAACTTTGATAAATTCTACATCATAGCATATTGAAGC-AG 100595
Qy 478 AATAACAAATGCTATACCTCAGGAATATTAATTCCAGATTTTACAGCATTTTAACTTTCT 537
Db 100596 AATAACAAATGCTATACCTCAGGAATATTAATTCCAGATTTTACAGCATTTTAACTTTCT 100655
Qy 538 TGATGAGAAAAAATAAATTTGTCAGTTATTAAACTATTT-GGATCCAACAGATGAAAGCA 596
Db 100656 TGATGAGAAAAAATAAATTTGTCAGTTATTAAACTATTTGGGATCCAACAGATGAAAGCA 100715
Qy 597 GAATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTT 656
Db 100716 GAATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTT 100775
Qy 657 GACACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTC 716
Db 100776 GACACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTC 100835
Qy 717 CTTTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTC 776
Db 100836 CTTTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTC 100895
Qy 777 AGCCTTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTTAACTGTAA 836
Db 100896 AG-CTTTTAAATATTTTCCCATTATAATTTTTG-GACCTCTAACTCTATTTTAACTGTAA 100953
Qy 837 ATATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTT 896
Db 100954 ATATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTT 101013
Qy 897 -AAAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTGC 955
Db 101014 AAAAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTTC 101073
Qy 956 ATGCC----ATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTG-AATGATTGAGT 1010
Db 101074 AATGCCATGAATCATCACAAATGGCAGGCTCATTGTGGTGGCATGCTGAAATGATTGAGT 101133
Qy 1011 GGGACTGTGCC----AGCTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACA 1066
Db 101134 GGGAACTGTGCCCAGCACTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACA 101193
Qy 1067 CACTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTG 1126
Db 101194 CACTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTG 101253
Qy 1127 CTCTTTTGTTTGGGCTTGGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACAC 1186
Db 101254 CTCTTTTGTTTGGGCTTGGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACAC 101313
Qy 1187 AATATTAAGATTATTGCAATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTT 1246
Db 101314 AATATTAAGATTATTGCAATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTT 101373
Qy 1247 TTGAGATTGCATTCCTTTCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTGTG 1306
Db 101374 TTGAGATTGCATTCCTTTCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTGTG 101433
Qy 1307 TGGTATGTTTACTGTAGTTGCTTAGAAGT-CCATTCTTTACCAATGCTCAAATGTGATTA 1365
Db 101434 TGGTATGTTTACTGTAGTTGCTTAGAAGTCCCATTCTTTACCAATGCTCAAATGTGATTA 101493
Qy 1366 AATTTG-TTTTCTTGTTAAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTTA-TT 1423
Db 101494 AATTTGTTTTTCTTGTTAAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTTATTT 101553
Qy 1424 TCAGTGATTTAACATTTCCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTTTCT 1483
Db 101554 TCAGTGATTTAACATTTCCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTTTCT 101613
Qy 1484 CCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACCAAC 1543
Db 101614 CCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACCAAC 101673
Qy 1544 ACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTGAAT 1603
Db 101674 ACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTGAAT 101733
Qy 1604 GACAGATACTTCAGGCTTTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAAGGT 1663
Db 101734 GACAGATACTTCAGGCTTTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAAGGT 101793
Qy 1664 TCAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAACTC 1723
Db 101794 TCAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAACTC 101853
Qy 1724 AATACTACCTTGTCTTTTAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCA 1783
Db 101854 AATACTACCTTGTCTTTTAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCA 101913
Qy 1784 ATACTTGCTTCATTTCTGGCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAA 1843
Db 101914 ATACTTGCTTCATTTCT-GCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAA 101972
Qy 1844 GGTGGCGAACATAAAACCTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGA 1903
Db 101973 GGTGGCG-ACATAAAACCTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGA 102031
Qy 1904 GATTAAGAGAGGCTAGCCGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTT 1963
Db 102032 GATTAAGAGAGGCTAGCCGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTT 102091
Qy 1964 GCAGGCTCTGGTAAGAAGTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACA 2023
Db 102092 GCAGGCTCTGGTAAGAAGTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACA 102151
Qy 2024 TATGCAAAATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGT 2083
Db 102152 TATGCAAAATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGT 102211
Qy 2084 CACTGAAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCA 2143
Db 102212 CACTGAAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCA 102271
Qy 2144 CACTGAGTTTTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACAC 2203
Db 102272 CACTGAGTTTTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACAC 102331
Qy 2204 GTCCTTTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAGTCACAGGCAAT 2263
Db 102332 GTCCTTTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAGTCACAGGCAAT 102391

```
Qy       2264 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 2323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102392 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 102451

Qy       2324 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 2383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102452 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 102511

Qy       2384 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 2443
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     102512 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAAGGAAGACACACTCGGTCTTCTT 102571

Qy       2444 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 2503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102572 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 102631

Qy       2504 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 2563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102632 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 102691

Qy       2564 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 2623
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102692 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 102751

Qy       2624 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGA 2682
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| |||
Db     102752 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGA 102811

Qy       2683 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102812 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 102871

Qy       2743 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 2802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102872 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 102931

Qy       2803 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102932 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 102991

Qy       2863 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 2922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102992 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 103051

Qy       2923 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 2982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     103052 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 103111

Qy       2983 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 3042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     103112 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 103171

Qy       3043 TCG 3045
             |||
Db     103172 TCG 103174
```

```
RESULT 3
AX194837/c
LOCUS       AX194837               3000 bp    DNA     linear   PAT 28-AUG-2001
DEFINITION  Sequence 307 from Patent WO0151659.
ACCESSION   AX194837
VERSION     AX194837.1  GI:15385484
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Chu,T., Blumenfeld,M. and Cohen,D.
  TITLE     Biallelic markers derived from genomic regions carrying genes
            involved in central nervous system disorders
  JOURNAL   Patent: WO 0151659-A 307 19-JUL-2001;
            GENSET (FR)
FEATURES             Location/Qualifiers
     source          1. .3000
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     primer_bind     1263. .1281
                     /note="downstream amplification primer"
     misc_feature    1481. .1500
                     /note="8-42-211.mis1, potential"
     misc_feature    1482. .1500
                     /note="8-42-211.mis2, complement"
     misc_feature    1489. .1513
                     /note="8-42-211 potential probe"
     variation       1501
                     /note="8-42-211 : polymorphic base C or G"
     primer_bind     1694. .1711
                     /note="upstream amplification primer, complement"
ORIGIN
```

Query Match 64.3%; Score 1959.2; DB 6; Length 3000;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2018; Conservative 1; Mismatches 1; Indels 6; Gaps 6;

```
Qy       1024 CTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACACACTGTCCTGTATTGTA 1083
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2900 CTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACACACTGTCCTGTATTGTA 2841

Qy       1084 ATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTGCTCTTTTGTTTGGGCTT 1143
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       2840 ATGCATTGGCCCAAGTGGATTCTTTTTGATGCTTTGGTGATTGCTCTTTTGTTTGGGCTT 2781

Qy       1144 GGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACACAATATTAAGATTATTGC 1203
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2780 GGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACACAATATTAAGATTATTGC 2721

Qy       1204 AATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTTTTGAGATTGCATTCCTT 1263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2720 AATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTTTTGAGATTGCATTCCTT 2661

Qy       1264 TCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTGTGTGGTATGTTTACTGTAG 1323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2660 TCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTGTGTGGTATGTTTACTGTAG 2601

Qy       1324 TTGCTTAGAAGT-CCATTCTTTACCAATGCTCAAATGTGATTAAATTTG-TTTTCTTGTT 1381
             |||||||||||| ||||||||||||||||||||||||||||||||||| |||||||||||
Db       2600 TTGCTTAGAAGTCCCATTCTTTACCAATGCTCAAATGTGATTAAATTTGTTTTTCTTGTT 2541

Qy       1382 AAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTTA-TTTCAGTGATTTAACATTT 1440
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       2540 AAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTTATTTTCAGTGATTTAACATTT 2481

Qy       1441 CCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTTTCTCCAACAAAAGGTAAATT 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2480 CCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTTTCTCCAACAAAAGGTAAATT 2421

Qy       1501 TATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACCAACACAGGTGAAAGTGTTAG 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2420 TATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACCAACACAGGTGAAAGTGTTAG 2361

Qy       1561 CCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTGAATGACAGATACTTCAGGCT 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2360 CCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTGAATGACAGATACTTCAGGCT 2301

Qy       1621 TTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAAGGTTCAGAGCAAAAGAGGGC 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2300 TTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAAGGTTCAGAGCAAAAGAGGGC 2241

Qy       1681 ACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAACTCAATACTACCTTGTCTTT 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2240 ACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAACTCAATACTACCTTGTCTTT 2181
```

```
Qy       1741 TAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCAATACTTGCTTCATTTCT 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2180 TAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCAATACTTGCTTCATTTCT 2121

Qy       1801 GGCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAAGGTGGCGAACATAAAAC 1860
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       2120 -GCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAAGGTGGCG-ACATAAAAC 2063

Qy       1861 CTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGAGATTAAGAGAGGCTAGC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2062 CTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGAGATTAAGAGAGGCTAGC 2003

Qy       1921 CGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTTGCAGGCTCTGGTAAGAA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2002 CGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTTGCAGGCTCTGGTAAGAA 1943

Qy       1981 GTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACATATGCAAAATATTTCCA 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1942 GTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACATATGCAAAATATTTCCA 1883

Qy       2041 TCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGTCACTGAAATTACAAGTG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1882 TCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGTCACTGAAATTACAAGTG 1823

Qy       2101 TAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCACACTGAGTTTTGTTCTT 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1822 TAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCACACTGAGTTTTGTTCTT 1763

Qy       2161 CATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACACGTCCTTTATAATTTCGT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1762 CATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACACGTCCTTTATAATTTCGT 1703

Qy       2221 TCTCTCCCGGTTCCCCAACGTTAAAAAAAAAAGTCACAGGCAATATTCTCCCTGAGGGAGT 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1702 TCTCTCCCGGTTCCCCAACGTTAAAAAAAAAAGTCACAGGCAATATTCTCCCTGAGGGAGT 1643

Qy       2281 AAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGTTGTTGTCGTCGTTGTTC 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1642 AAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGTTGTTGTCGTCGTTGTTC 1583

Qy       2341 GTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAATGGCGCGAGA 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1582 GTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAATGGCGCGAGA 1523

Qy       2401 ACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATCAATTAGCAATAA 2460
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db       1522 ACGGAGGTAGCTTTTTAAAAASGAAGACACACTCGGTCTTCTTCCATCAATTAGCAATAA 1463

Qy       2461 TTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTATGCTTCCTTTTCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1462 TTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTATGCTTCCTTTTCT 1403

Qy       2521 CATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGCAGATAATATGAG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1402 CATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGCAGATAATATGAG 1343

Qy       2581 GCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTTTTGCTGGGTTGG 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1342 GCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTTTTGCTGGGTTGG 1283

Qy       2641 AGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGATCTGGTGTGTGTAAGTG 2699
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       1282 AGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGATCTGGTGTGTGTAAGTG 1223

Qy       2700 AGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTGAGCTAGAGGGAG 2759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1222 AGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTGAGCTAGAGGGAG 1163

Qy       2760 AGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAAGTATTAGGAGGG 2819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1162 AGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAAGTATTAGGAGGG 1103

Qy       2820 GAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGACTCACAGAGGGATAAATA 2879
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1102 GAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGACTCACAGAGGGATAAATA 1043

Qy       2880 AAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGGTGGGGAGAAGGGGGACG 2939
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1042 AAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGGTGGGGAGAAGGGGGACG 983

Qy       2940 AAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGGAAGGAAGGAAAT 2999
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        982 AAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGGAAGGAAGGAAAT 923

Qy       3000 AGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 3045
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        922 AGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 877
```

RESULT 4
AX194838
LOCUS AX194838 3001 bp DNA linear PAT 28-AUG-2001
DEFINITION Sequence 308 from Patent WO0151659.
ACCESSION AX194838
VERSION AX194838.1 GI:15385485
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chu,T., Blumenfeld,M. and Cohen,D.
TITLE Biallelic markers derived from genomic regions carrying genes involved in central nervous system disorders
JOURNAL Patent: WO 0151659-A 308 19-JUL-2001;
GENSET (FR)

FEATURES	Location/Qualifiers
source	1. .3001
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"
primer_bind	1114. .1133
	/note="upstream amplification primer"
misc_feature	1481. .1500
	/note="8-45-389.mis1, potential"
misc_feature	1489. .1513
	/note="8-45-389 potential probe"
variation	1501
	/note="8-45-389 : polymorphic base A or G"
misc_feature	1502. .1520
	/note="8-45-389.mis2, complement"
primer_bind	1516. .1533
	/note="downstream amplification primer, complement"

ORIGIN

Query Match 32.2%; Score 979.4; DB 6; Length 3001;
Best Local Similarity 98.7%; Pred. No. 3.2e-161;
Matches 1005; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

```
Qy       2029 AAAATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGTCACTG 2088
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 AANATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGTCACTG 68

Qy       2089 AAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCACACTG 2148
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db         69 AAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCACANTG 128

Qy       2149 AGTTTTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACACGTCCT 2208
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         129 AGTTNTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACACGTCCT 188

Qy       2209 TTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAGTCACAGGCAATATTCT 2268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         189 TTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAGTCACAGGCAATATTCT 248

Qy       2269 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGTTGTTG 2328
```

```
Db        249 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGNAGGTACCGTTTTGTTGTTGTTG 308

Qy       2329 TCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGC 2388

Db        309 TCGTCGNTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGC 368

Qy       2389 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATC 2448

Db        369 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATC 428

Qy       2449 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 2508

Db        429 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 488

Qy       2509 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 2568

Db        489 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 548

Qy       2569 AGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 2628

Db        549 AGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 608

Qy       2629 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGATCTGG 2687

Db        609 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGACCTGG 668

Qy       2688 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 2747

Db        669 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 728

Qy       2748 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 2807

Db        729 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 788

Qy       2808 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGACTCACA 2867

Db        789 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCGCTTTCGAAGCGACTCACA 847

Qy       2868 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGGTGGGG 2927

Db        848 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTAAAGGGAAGGCAGGTGGGG 907

Qy       2928 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 2987

Db        908 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 967

Qy       2988 AAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 3045

Db        968 AAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 1025
```

RESULT 5
HS5HT1A

```
LOCUS       HS5HT1A                 1204 bp    DNA     linear   PRI 12-AUG-1996
DEFINITION  H.sapiens 5HT1A receptor region.
ACCESSION   Z11168 X60451
VERSION     Z11168.1  GI:1033027
KEYWORDS    promoter; serotonin receptor protein.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1204)
  AUTHORS   Parks,C.L., Chang,L.S. and Shenk,T.
  TITLE     A polymerase chain reaction mediated by a single primer: cloning of
            genomic sequences adjacent to a serotonin receptor protein coding
            region
  JOURNAL   Nucleic Acids Res. 19 (25), 7155-7160 (1991)
  MEDLINE   92115564
   PUBMED   1766875
REFERENCE   2  (bases 1 to 1204)
  AUTHORS   Parks,C.L. and Shenk,T.
  TITLE     The serotonin 1a receptor gene contains a TATA-less promoter that
            responds to MAZ and Sp1
  JOURNAL   J. Biol. Chem. 271 (8), 4417-4430 (1996)
  MEDLINE   96224025
   PUBMED   8626793
REFERENCE   3  (bases 1 to 1204)
  AUTHORS   Parks,C.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-SEP-1991) C.L. Parks, Lewis Thomas Laboratory,
            Princeton University, Dept of Mol Biology, Princeton, New Jersey
            08544-1014, USA
COMMENT     On Oct 22, 1995 this sequence version replaced gi:23893.
            Overlaps cDNA sequences; Kobilka et al., Nature 329:75-79(1987),
            Fargin et al., Nature 335:35-360(1988).
FEATURES             Location/Qualifiers
     source          1. .1204
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /cell_line="human 293 cells"
                     /clone_lib="DNA prepared by PCR"
     promoter        1. .1176
                     /note="promoter activity in transient expression assay"
     mRNA            393. .>1204
     repeat_region   661. .745
                     /note="polypurine stretch"
     mRNA            1159. .>1204
     CDS             1177. .>1204
                     /note="initiation ATG codon"
                     /codon_start=1
                     /product="5-HT1a receptor"
                     /protein_id="CAA77560.1"
                     /db_xref="GI:1033028"
                     /translation="MDVLSPGQG"
ORIGIN
```

Query Match 24.6%; Score 747.8; DB 9; Length 1204;
Best Local Similarity 98.9%; Pred. No. 9.9e-121;
Matches 774; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

```
Qy       2264 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 2323

Db          1 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAGCGGAGGTACCGTTTTGTTGT 60

Qy       2324 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 2383

Db         61 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 120

Qy       2384 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 2443

Db        121 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 180

Qy       2444 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 2503

Db        181 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 240

Qy       2504 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 2563

Db        241 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 300

Qy       2564 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 2623

Db        301 TTGGCAGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 360

Qy       2624 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGA 2682

Db        361 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGA 420

Qy       2683 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 2742

Db        421 CCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 480

Qy       2743 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 2802
```

```
Db        481 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 540

Qy       2803 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 2862
             ||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||
Db        541 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCGCTTTCGAAGCGAC 599

Qy       2863 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 2922
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        600 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTAAAGGGAAGGCAGG 659

Qy       2923 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 2982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 719

Qy       2983 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 3042
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        720 GAGGGAAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 779

Qy       3043 TCG 3045
              |||
Db        780 TCG 782


RESULT 6
AX194847
LOCUS       AX194847               2512 bp    DNA     linear   PAT 28-AUG-2001
DEFINITION  Sequence 317 from Patent WO0151659.
ACCESSION   AX194847
VERSION     AX194847.1  GI:15385494
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Chu,T., Blumenfeld,M. and Cohen,D.
  TITLE     Biallelic markers derived from genomic regions carrying genes
            involved in central nervous system disorders
  JOURNAL   Patent: WO 0151659-A 317 19-JUL-2001;
            GENSET (FR)
FEATURES             Location/Qualifiers
     source          1. .2512
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     primer_bind     1171. .1189
                     /note="upstream amplification primer"
     misc_feature    1482. .1500
                     /note="99-5069-331.mis1"
     misc_feature    1489. .1513
                     /note="99-5069-331 potential probe"
     variation       1501
                     /note="99-5069-331 : polymorphic base C or T"
     misc_feature    1502. .1521
                     /note="99-5069-331.mis2, potential complement"
     primer_bind     1702. .1719
                     /note="downstream amplification primer, complement"
ORIGIN

  Query Match             17.2%;  Score 522.4;  DB 6;  Length 2512;
  Best Local Similarity   98.6%;  Pred. No. 1.9e-81;
  Matches  548;  Conservative    0;  Mismatches    6;  Indels    2;  Gaps    2;

Qy       2491 CCCATTCAGGCTCCCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGAC 2550
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCATTCAGGCTCCCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGAC 60

Qy       2551 ACGATTTAAGAATTTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAA 2610
             |||||||||||||||||||||||||||||   ||||||||||||||||||||||||||||
Db         61 ACGATTTAAGAATTTGGCAGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAA 120

Qy       2611 GTTTTTCCAACCCCAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTC 2670
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTTTTTCCAACCCCAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTC 180

Qy       2671 TGACCGGCA-GGATCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGAC 2729
             |||||||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGACCGGCAGGGACCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGAC 240

Qy       2730 TCGAATGCAAAGACGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCA 2789
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCGAATGCAAAGACGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCA 300

Qy       2790 CTCCTCGGGGTTGGGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGG 2849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||| || | |
Db        301 CTCCTCGGGGTTGGGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCG 359

Qy       2850 CTTTCGAAGCGACTCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGA 2909
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        360 CTTTCGAAGCGACTCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTA 419

Qy       2910 AAGGGAAGGCAGGTGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGG 2969
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 AAGGGAAGGCAGGTGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGG 479

Qy       2970 AGAGAGGGGGAGAGAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGG 3029
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        480 AGAGAGGGGGAGAGAGGGAAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGG 539

Qy       3030 AGGATGGGGCTTCTCG 3045
              ||||||||||||||||
Db        540 AGGATGGGGCTTCTCG 555


RESULT 7
HS5HT1AR5
LOCUS       HS5HT1AR5              699 bp    DNA     linear   PRI 12-JUL-1994
DEFINITION  H.sapiens DNA 5' to 5HT1A receptor gene.
ACCESSION   X53826
VERSION     X53826.1  GI:510982
KEYWORDS    palindrome.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 699)
  AUTHORS   Potts,W.J., Morten,J., Elvin,P., Deakin,W., Braude,W.M.,
            Waring,H.L. and Jenner,D.E.
  TITLE     Sequence 5' to the gene for human 5HT1A receptor obtained by
            'vectorette' walking
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 699)
  AUTHORS   Jenner,D.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUL-1990) D.E. Jenner, ICI PHARMACEUTICALS, ALDERLEY
            PARK, MACCLESFIELD, CHESHIRE, SK10 4TG, UK
FEATURES             Location/Qualifiers
     source          1. .699
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5q11-5q13"
     repeat_unit     606. .621
                     /note="palindrome"
                     /rpt_type=INVERTED
     repeat_unit     681. .696
                     /note="palindrome (to above)"
                     /rpt_type=INVERTED
ORIGIN

  Query Match             14.4%;  Score 438;  DB 9;  Length 699;
  Best Local Similarity   99.8%;  Pred. No. 1.3e-66;
  Matches  449;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;
```

```
Qy       2596 AATTCCCTCCCCCAAGTTTTTCCAACCCCAGTTTTGCTGGGTTGGAGGCGGAGTTTATTT 2655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 AATTCCCTCCCCCAAGTTTTTCCAACCCCAGTTTTGCTGGGTTGGAGGCGGAGTTTATTT 82

Qy       2656 GTTACAACCTTGGTCTGACCGGCAGGATCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGT 2715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 GTTACAACCTTGGTCTGACCGGCAGGATCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGT 142

Qy       2716 TGACAAAAAGAGACTCGAATGCAAAGACGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCC 2775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 TGACAAAAAGAGACTCGAATGCAAAGACGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCC 202

Qy       2776 AGAGGAAAGAGGCACTCCTCGGGGTTGGGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGA 2835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 AGAGGAAAGAGGCACTCCTCGGGGTTGGGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGA 262

Qy       2836 GGGAAGGAGCCTGGCTTTCGAAGCGACTCACAGAGGGATAAATAAAGGGAAGTGAGGAGG 2895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 GGGAAGGAGCCTGGCTTTCGAAGCGACTCACAGAGGGATAAATAAAGGGAAGTGAGGAGG 322

Qy       2896 AAGAGGGAGACTGAAAGGGAAGGCAGGTGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGA 2955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 AAGAGGGAGACTGAAAGGGAAGGCAGGTGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGA 382

Qy       2956 GAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCA 3015
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        383 GAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGGAAGGAAGGAAATAGGGAGAGGA-GGTCA 441

Qy       3016 CAGAGTGACCGTGGAGGATGGGGCTTCTCG 3045
              ||||||||||||||||||||||||||||||
Db        442 CAGAGTGACCGTGGAGGATGGGGCTTCTCG 471

RESULT 8
AF087675
LOCUS       AF087675                3635 bp    DNA     linear   ROD 14-FEB-2002
DEFINITION  Rattus norvegicus serotonin 1A receptor gene, partial cds.
ACCESSION   AF087675
VERSION     AF087675.1  GI:3603449
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 3635)
  AUTHORS   Meijer,O.C., Williamson,A., Dallman,M.F. and Pearce,D.
  TITLE     Transcriptional repression of the 5-HT1A receptor promoter by
            corticosterone via mineralocorticoid receptors depends on the
            cellular context
  JOURNAL   J. Neuroendocrinol. 12 (3), 245-254 (2000)
  MEDLINE   20183470
   PUBMED   10718920
REFERENCE   2  (bases 1 to 3635)
  AUTHORS   Meijer,O.C., Dallman,M.F. and Pearce,D.
  TITLE     The rat 5-HT1A receptor promoter: activation by NGFI-C, sp1 and
            NF-kB, and context dependent repression by corticosteroids
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 3635)
  AUTHORS   Meijer,O.C. and Pearce,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-AUG-1998) Medicine, UCSF, 1001 Portero Avenue, San
            Francisco, CA 94110, USA
FEATURES             Location/Qualifiers
     source          1..3635
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
     CDS             3610..>3635
                     /note="5-HT1A receptor"
                     /codon_start=1
                     /product="serotonin 1A receptor"
                     /protein_id="AAC35855.1"
                     /db_xref="GI:3603450"
                     /translation="MDVFSFGQG"
ORIGIN

  Query Match             5.4%;  Score 163.6;  DB 10;  Length 3635;
  Best Local Similarity   53.9%;  Pred. No. 7.4e-19;
  Matches  893;  Conservative    2;  Mismatches  626;  Indels  135;  Gaps   22;

Qy        136 TTAAGAGTTAAACATAGAAGAATTGGCTAAG-TGAAAATGAATGAAACGCAATATCATTC 194
Db        196 TTAAGCGTTAAACATAAACCGATTGACTAAGCATGGCATGAAGAAATGCAATACTAGTC 255

Qy        195 TGCATATATCATTTATTATATATCACAGTATTATTAGTTTTAAAAGTTAAACATAAATAT 254
Db        256 TGATTTTATTATATA------ATCACAGTATTATAAGTTTAATAACAACAAAAAATCATA 309

Qy        255 CTATTATGYCATTGSACGAYTAGGYSAACCTARTCRGTGC-TGCGAATACTTTCGATACT 313
Db        310 AAATTATTAGTACCTAAAAGATGCAGAACTATTACACTTCATGTGCATTCTTCCTAGCAT 369

Qy        314 TCTGTTTCCCTCCTAGTATTCATAAGTGTGCCTTTGAAAACGTTTTAAATTGT------A 367
Db        370 CCAGAAATACAGAAAGGTGTCTACAATCTACATTTTAAGAGAGCAAAAAATTTTTTAAAA 429

Qy        368 AGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACTTGAATTACTTT 427
Db        430 AGAAAGCAAATGCTTGATATATCATGGGCATTAATTTAAGGATGAATATGAACTACTTTG 489

Qy        428 GGATTTTGAAAA------------------ACTTTGATAAATTCTACATCATAGCATAT 468
Db        490 GAATTTCCTAAACATGAATTAAACACGCTACACTCTGCTATATCAAACATGATACCCTCT 549

Qy        469 TGAAGCAAGAATAACAAATGCTATACCTCAGGAATATTAATTCCAGAT--------TTTA 520
Db        550 ATCAAATTAAAATATAAAGACAAATCCAAGTTATTAAGATATAAACATAAACTGCACTCA 609

Qy        521 CAGCATTTTAACTTTCTTGATGAGAAAAAATAAATTTGTCAGTTATTAAACTATTTGGAT 580
Db        610 AAGCCTTCTAATTTTATTGAAAAGGAAAAGTAAGCTTGTTAGCTATTAAACCAATGAGGA 669

Qy        581 CCAACAGATGAAAGCAGAATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATA 640
Db        670 TCCAGTATACAAAGGACAATT---ACTGACTGACACGTTTCCTGATTTGTGACTTACGTA 726

Qy        641 TTTACATGTGTTGTTTGACACAATT-----CTTAATTATGTTCTTGATATGCATATATTT 695
Db        727 TTTACATATGCTAAGTCTGACACATTAGCATTCACATTGGTTCTTGAGGAATACATATTT 786

Qy        696 GCTTCTTAAATTTTAAGTTTCCTTTATTTTACTTTGTTTATAGTCTCAACTATAATTTCA 755
Db        787 GCTTTTTGAATGTACATGTGTTTCTTCTT---TTACTCTGTAGTCTTGGCCACATTTTTA 843

Qy        756 AAGTTTAATTTTAGATAATTCAGCCTTTTAAATATTTTCCCATTATAATTTTTGTGACCT 815
Db        844 AAGTTTTCTCCTTGGAAATTTTAGATGTTAAGCAT----CCATTATAATCTCTGC-AGTT 898

Qy        816 CTAACTCTATTTTAACTGTAAATATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAA 875
Db        899 GGAACTATATTTTAA---TAAACATTTTTAACTGTTTCGTGACGGATTTCACTAGAACAA 955

Qy        876 ATAGATACCTTCACAAATCTTAAAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGT 935
Db        956 CTTACATGCTCACATGTTTTAAAAGGCGTTCCTCAGAGCCTGT-------CTTGCTGTGC 1008

Qy        936 ATACTTATCTCTTTCTTTGCATGCCATGATCATCACAATGCATGGCTCATGTGGTGGCAT 995
Db       1009 ACAGTTGCTGCTTTGTTTGGTGTTATGAATCCGCACACATGACAGCTTCAGGGCTTCCCT 1068

Qy        996 GCTGAATGATTGAGTGGGACTGTGCCAGCTGAACTATAAAAAAAAAAAACAAACAAAACC 1055
Db       1069 ATTCTCAAATGATCGAGCCGGGACTGTGC------------------CCACATCAGTCC 1109
```

```
Qy       1056 TTATCCAAACACACTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGC 1115
Db       1110 ATCTTCAGGTACACAGCCCTGGGTTAT-ATATTTCTGCCCAACTGGACTTTCTTTGGTGC 1168

Qy       1116 TTTGGTGATTGCTCTTTTGTTTGGGCTTGGAGAATTCAGAGCTATGAAATTCAGAGCTCA 1175
Db       1169 TTTAGTGATTACTCTTTTCTTT-GGTTTTAAGGTTTCAAAACTATGACATCCAAAGCTCA 1227

Qy       1176 GATTTGAACACAATATTAAGATTATTGCAATCTGTAGTGAATCTGTTCATGTTATCCAGT 1235
Db       1228 GATGTCAACACAACAGTAAGACAATTGCAGTTTGTAGCGGGTCCTTGCCTGTTGTCTAGA 1287

Qy       1236 GTCAACTGCTTTTGAGATTGCATT-----CCTTTCACCTCAGGCATGCAATCAGGATGTA 1290
Db       1288 ATTGGCTGCTGTGAGAATTGCATTATGACTTCTTCAGTTCAGGTATGTGATCAGATTCTC 1347

Qy       1291 TAAGTGAAATGTTGTGTG----GTATGTTTACTGTAGTTGCTTAGAAGTCCATTCTTTAC 1346
Db       1348 TAAGCAGATCTTTTTGTGTGCTGTTTTTTTATTCTGATTAATTAGAAGTGCAATTCTT-- 1405

Qy       1347 CAATGCTCAAATGTGATTAAATTTGTTTTCTTGTTAAAGGAAACAGCTTAGAACAAACCC 1406
Db       1406 -TATCCTCAGTTGGGATTGCTTTTGTTTTTCCAAAAATACAATCAGCTTAAAACAAATTC 1464

Qy       1407 TTGTAAGTATCTTTATTTCAGTGATTTAACATTTCCAAATGTTAAATCATTTGGAAAATG 1466
Db       1465 CTAAATGTCTCTAATTTTCAGTATTAAGTCATTTCCTTATGA-----------AAAAGG 1512

Qy       1467 CAATACTATTCGTTTCTCCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTCAGGTTAT 1526
Db       1513 AAATAATGCCCTCTTCCCCCAATAAAAC-------ATGTGAGTTCTATAGCACAAACTGT 1565

Qy       1527 GACAGCACAAAACCAACACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGGCATTCCC 1586
Db       1566 GACACCACAGCACAAATAGCCGAAAATGTTGTTTGCCATGTTTCCTA------------- 1612

Qy       1587 AGTTAGAACTTGTGAATGACAGATACTTCAGGCTTTCGAAGGAAGCTAAAACATATAATA 1646
Db       1613 ---TAGAAGTTGTGAATGATAGATGTCATAAGCTTTCAAAGGGAGCTAGAACATACGCCA 1669

Qy       1647 GGCCTGATATATAAGGTTCAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATA 1706
Db       1670 GGCTTGGTACAAACACCCAAGAAAAAGAAAAGAAAGAGAAAAGAAATGATTGAAGAAAAA 1729

Qy       1707 GGAAGGAGACAAAACTCAATACTACCTTGTCTTTTA 1742
Db       1730 ATTTAAAGTGACATTTAATGGATTCTATTTATTTTA 1765
```

RESULT 9
AC114573

```
LOCUS       AC114573               203941 bp    DNA     linear   HTG 17-DEC-2003
DEFINITION  Mus musculus chromosome 13 clone RP23-146M15 map 13, *** SEQUENCING
            IN PROGRESS ***, 5 unordered pieces.
ACCESSION   AC114573
VERSION     AC114573.5  GI:39979456
KEYWORDS    HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 203941)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 13, clone RP23-146M15
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 203941)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 203941)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-DEC-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Dec 17, 2003 this sequence version replaced gi:28882292.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L23431
                Center clone name: 146_M_15
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    41236: contig of 41236 bp in length
            *    41237    41336: gap of 100 bp
            *    41337    53790: contig of 12454 bp in length
            *    53791    53890: gap of 100 bp
            *    53891    58632: contig of 4742 bp in length
            *    58633    58732: gap of 100 bp
            *    58733   109660: contig of 50928 bp in length
            *   109661   109760: gap of 100 bp
            *   109761   203941: contig of 94181 bp in length.
FEATURES             Location/Qualifiers
```

```
    source          1. .203941
                    /organism="Mus musculus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10090"
                    /chromosome="13"
                    /map="13"
                    /clone="RP23-146M15"
                    /clone_lib="RPCI-23 Female Mouse BAC"
ORIGIN

  Query Match             5.3%;  Score 162.6;  DB 2;  Length 203941;
  Best Local Similarity   49.9%;  Pred. No. 5.4e-19;
  Matches 1390;  Conservative    0;  Mismatches 1199;  Indels  198;  Gaps   31;

Qy        357 TTTAAATTGTAAGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACT 416
Db      33347 TCTACATTTTAAGAAAGCAAATGTTTGATATATCATGTACATTAATT-TGAGGATGAATA 33405

Qy        417 TGAATTACTTTGGATTTTGAAAAACTTTGAT---AAATTCTACATCATAGCATATTGAAG 473
Db      33406 TGAGCTACTTGGGAGTTTCCTAAACTTGAATGTGACACGGTACACTCTGTCATATT-AAG 33464

Qy        474 CAAGAATAACAAATGCTATACCTCAGGAATATTAATTCCAGATTTTACAGCATTTTAACT 533
Db      33465 ATATAAAATCAAATCCCAGACTTTATAAACATAAGCTTCAGACTCAA-AGCCTCTTAATT 33523

Qy        534 TTCTTGATGAGAAAAAATAAATTTGTCAGTTATTAAACTATTTGGATCCAACAGATGAAA 593
Db      33524 TTACTGAAAAGGAAAAATGAGTTTGTTAGCTATTAACCGAATTCGGATCCAGTACACAAA 33583

Qy        594 GCAGAATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTG 653
Db      33584 GAACAATT---GCTGACTGACATGTTTCCTGATTAGTGACTTATATATTTGCATATGCTA 33640

Qy        654 TTTGACAC-----AATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTT 708
Db      33641 CGTCTGACCCATTAATATTCAAGTTGGATCTAGATATGAACATATTTGCTTTTGAATAT 33700

Qy        709 TAAGTTTCCTTTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTA 768
Db      33701 AAATGTTTCTTATTTTTACTCTG----TAGTCTTAGCTACATTCTTAAAATGTTGTCCTC 33756

Qy        769 GATAATTCAGCCTTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTT 828
Db      33757 GGAAATTTTAGATGTTAAGTGT----CCCTTATAATTTCTG-CAGTTGCAACTATATTTT 33811

Qy        829 AACTGTAAATATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCA 888
Db      33812 AATAAACATTTTAGCTCTGTTTTG--TGAGAGATTTTATCAGAATATCTTACATGTTTAC 33869

Qy        889 CAAATCTTAAAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTT 948
Db      33870 ATGCCTTTAAAAGCATTTCTCAAAGTCTGTC-------TTGCTGTGCATACTTGATGCTA 33922

Qy        949 TCTTTGCATGCCATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTGAATGATTGA 1008
Db      33923 ATTTTTTGTGTTATGA-----------------------ATGGCAGCTTCAGTGTTTTC 33958

Qy       1009 GTGGGACTGTGCCAGCTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACACA 1068
Db      33959 CTATCCTGAAATGATGGGGCCAATCCATCTTTGGGCACATAGCCCTGTATTGTATATTGT 34018

Qy       1069 CTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTGCT 1128
Db      34019 ATATGTCTGTATTGTATATGTCTGACCAACTGGACTTTCTTTGATCCTTTTAGTGATTAC 34078

Qy       1129 CTTTTGTTTGGGCTTGGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACACAA 1188
Db      34079 TCTTTCCTTTGGTTTTAAGATTTCAAAACTATGACATTCAAAGCTAAGATTTGAACACAA 34138

Qy       1189 T---------------ATTAAGATTATTGCAATCTGTAGTGAATCTGTTCATGTTATCCA 1233
Db      34139 TTCGAGTTCTGTACACAGTAAGACAATTGTAATTTGTAGTGAGTCCTTGCCTGTTATCTA 34198

Qy       1234 GTGTCAACTGCTTTTGAGATTGCAT-----TCCTTTCACCTCAGGCATGCAATCAGGATG 1288
Db      34199 GAATTAGCAGCTCTGAGAATTGCATCATGACTATGTCACTTCAAGAATGTAATCAGAATC 34258

Qy       1289 TATAAGTGAAA--TGTTGTGTGGTATGTTTACTGTAGTTGCTTAGAAGTCCATTCTTTAC 1346
Db      34259 TCTAAGCACAACTTTTTGTGTGCTGTTTTTATTCTGATTAATTAGAAGTGC---CATTCT 34315

Qy       1347 CAATGCTCAAATGTGATTAAATTTGTTTTCTTGTTAAAGGAAACAGCTTAGAACAAACCC 1406
Db      34316 TTGTGTTCAATTGGGATTGATTTTTTTTCAAAAATAAAGG---CAGCTTAAAACAAACTC 34372

Qy       1407 TTGTAAGTATCTTTATTTCAGTGATTTAACATTTCCAAATGTTAAATCATTTGGAAAATG 1466
Db      34373 TTAGATGTCTCTAATTTT------------------AGTATCAAATTATTTCTTTATGC 34413

Qy       1467 CAATACTATTCGTTTCTCCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTCAGGTTAT 1526
Db      34414 AAAAGGTAATACTGTCACCATCCCATAATAAAACACTTGAGTTCTAATAACACAAGCTGC 34473

Qy       1527 GACAGCACAAAACCAACACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGGCATTCCC 1586
Db      34474 TGCACCACAACATAAGTAGCTGAAAGTGTTTTCTGCCA---------------CTGTTCC 34518

Qy       1587 AGTTAGAACTTGTGAATGACAGATACTTCAGGCTTTCGAAGGAAGCTAAAACATATAATA 1646
Db      34519 TATAGATTTTTGTGAATGATAGATGTCATAAGCTTTCAAAGGGAGCTAGAAAATATACTT 34578

Qy       1647 GGCCTGATATATAAGGTTCAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAAT- 1705
Db      34579 GGCTTGGTATAAAAACCCAAGGGAAAGAAAGAGAAAAGAAATGGTTGAAGAAAAAATTTA 34638

Qy       1706 AGGAAGGAGACAAAACTCAATACTACCTTGTCTTTTAATAACTGTCTTCCTCTTTCTAAA 1765
Db      34639 AAGTGACATTTAAGGATTTTATTTATTTTATAATATTCTAGCCTTCCTAATGCTGTGATC 34698

Qy       1766 AGTTGTTGTATTTCCTCAATACTTGCTTCATTTCTGGCATAAGGGTTTCCAGATGGCACT 1825
Db      34699 CTTTAGCACAGTTCCTCATGTTGTGCAGTGCTCCCCAACCATGATATTATTTTCGTTGCT 34758

Qy       1826 CTAAAACATTTGCCAGAAGGTGGCGAACATAAAACCTCATTGCTTAGAACTGTCCCAGGT 1885
Db      34759 ACTTCATAAGTAGTAATTTGCTACTACAAATATTTTTGGAAGTATAGGTTCACAGATTGG 34818

Qy       1886 GCTGAACCCAGTTTCTGAGATTAAGAGAGGCTAGCCGGCTAGCGAACCGGGATTCC---- 1941
Db      34819 GTTTTGATACATTGGTTGTAAATAGATATCGTAGAGTCTTTGGGTTTCCTTAATACTTAC 34878

Qy       1942 --ACCAAGTTTCCCCCAGAGGTTTGCAGGCTCTGGTAAGAAGTGCAAAAGGCCATGTGAA 1999
Db      34879 TTAATTTTTTAATGACAAATATTTCTAAATGGCGCTCTGAAGCACTTGCCGGAAGGTGGC 34938

Qy       2000 ATGCCAGGC-TTCACTTAGAACACATATGCAAAATATTTCCATCCCTGAATTTACTAGCC 2058
Db      34939 GGCATAAGCAAGCTCTTACTACACAGAGCTGTCCCGGGTGCTGAACCCAATTTCAGAGAG 34998

Qy       2059 ACAAAGCTATGGGAAGTGGCAGTGTCACTGAAATTACAAGTGTAGTAGTGATGGAAAAGT 2118
Db      34999 ACTCGCAGTTGCTGGAGACAGGGATGCCACTAAGTTCCCTGGAGGGCTGCAGACTCCAGT 35058

Qy       2119 GTGTGTGTGTTTAGAATATATATCACACTGAGTTTTGTTCTTCATTTCGAGATGCAGTTG 2178
Db      35059 GAGAGAGACATAAGTCCAAAGACCAGTAGAAAAAGAAAGCCATTGTTAACAATATACAGG 35118

Qy       2179 TTTACCTCTCCTTGTCCTTTGACACGTCCTTTATAATTT--------CGTTCTCTCCCGG 2230
Db      35119 CAAACAATTTCTTTCCATGAATTTATTTGCTTCAAATCTATGGGAAATGTTAATATCATT 35178

Qy       2231 TTCCCCAACGTTAAAAAAAAAGTCACAGGCAATATTCTCCCTGAGGGAGTAAGGCTGGAC 2290
Db      35179 GATATGAAGAGAGTGACTGAAGAGTGTGTGTTTGGAATACACACTTAAGTGAATTTTTAA 35238
```

```
Qy       2291 TGTTAGATGATAACGGAGGTACCGTTTTGTTGTTGTTGTCGTCGTTGTTCGTTTGTTTT- 2349
             ||||   |   |   ||      ||  | |  |||  ||      ||| | ||    ||
Db      35239 TGTTTTCTTTGAGATTAGACTGTTTTCAGATCCTGTCCTCCAACCTGTCCTTTCTAATTT 35298

Qy       2350 --------TGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAATGGCGCGAGAA 2401
                     |  | | |||   |           | |||||  ||||  | |      |
Db      35299 CCCTCCAATACACAAGGATAATATGAAAAGGTCACAGGCACCAGTGATAGGTTAATATGG 35358

Qy       2402 CGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATCAATTAGCAATAAT 2461
              | | | || |||   ||  |    | || |    |   | | | |    | ||
Db      35359 TAAATGAACCTGTTTCTTAAGAAGCTGAGACCCACATTCTCTACCTGATAATGTAGAGGG 35418

Qy       2462 TGGGAGACTGACCCAGGACTGTTCACCTTCCCATTC-AGGCTCCCTATGCTTCCTTTTCT 2520
              | |    ||   || |   ||| |||||||| |      || ||    ||| ||
Db      35419 TAGAGTTTAGAAGGTGGTCGACTCAGCTTCCCATCCTTCAGAGTCTCTGAGGGTTTTCCT 35478

Qy       2521 CATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGCAGATAATATGAG 2580
              | |  | || ||||| ||| ||  ||| | | |||||||||  || |||| |  | | ||
Db      35479 CGTGCCTGATAGCCACACTGTGAACCTGGCTCTATTTAAGAAGCTGACAGAAATAAAGGG 35538

Qy       2581 GCAAGGAG-------------------------------TAGTTGGAATTCCCTCCCCC 2608
              |    | |                              || ||    ||  |
Db      35539 GATCTGCGGTGGGGAAATATATCTTTCTATCTCTTTCTTTTATTTTATTTTTTCCTTTTG 35598

Qy       2609 AAGTTTTTCCAACCCCAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGG 2668
              |||||||  | | |||||   ||   | |  ||  | ||||| ||| |||
Db      35599 TCTTTTTTCCCCCTCTAGTTTAATAGGTAAGAGGCGGGGTTTAGTATGTTAAAACATTGA 35658

Qy       2669 TCTGACCGGCAGGATCTGGTGTGTGT---AAGTGAGTTCTGAGTCTCTGTTGACAAAAAG 2725
              ||||||  ||||    ||  | |   |||||| |||||||  |    ||
Db      35659 TCTGACTAGCAGATCAGAGTAGATATATGAAGTGAATTCTGAGTGTGGTGTG-------- 35710

Qy       2726 AGACTCGAATGCAAAGACGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGA 2785
              ||||  ||||| |||||| |   |||| ||||||||||||||||||||||||| ||| ||
Db      35711 -GACTGGAATGTAAAGACTCCTGGCTACAGGGAGAGGAGGGCGGGGACCCAGAAGAAGGA 35769

Qy       2786 GGCACTCCTCGGGGTTGGGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGC 2845
              |||||||          |  |    || || | ||||||  |||||||| | ||||
Db      35770 GGCACTCGGGACTAGGGAAGTGTTGGATGGCGGTGGGGGTTAACGTGGGAGGAAGGGAGA 35829

Qy       2846 CTGGCTTTCGAAGCGACTCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGA 2905
              || ||| ||  |||||| ||| |||  |||| |||| || ||||| ||||| |||| |
Db      35830 CTAGCTGTCTGAGCGACACAC--AGGACTAAAGAAAGAGAGGTGAGAAGGAA-AGGGGAA 35886

Qy       2906 CTGAAAGGGAAGGCAGGTGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAG 2965
              ||| ||||||||||  |||||  ||||  ||  ||| || || |  |  | |||||||
Db      35887 CTGGAAGGGAAGGCCAGTGGGAGGAAGCCGGGAGAAGGAAGCTGGGAAAAGAGAAGCAGA 35946

Qy       2966 GAGGAGAGAGGGGGAGAGAGGGA-------AGGAAGGAAATAGGGAGAGGAGGGTCACAG 3018
                |||||  |  ||||||| |||       ||  | | ||   |||||  |||||||  |
Db      35947 CTGGAGAAGGCAGGAGAGATGGAGAGACAGAGAGAAGGAACTAGGAGATGAGGGTCATCG 36006

Qy       3019 AGTGACCGTGGAGGATGGGGCTTCTCG 3045
              |||||||  ||||||||||| |||| |
Db      36007 AGTGACCCGGGAGGATGGGGTTTCTGG 36033
```

RESULT 10
AC130951
LOCUS AC130951 259367 bp DNA linear HTG 15-NOV-2002
DEFINITION Rattus norvegicus clone CH230-28H16, *** SEQUENCING IN PROGRESS ***, 7 unordered pieces.
ACCESSION AC130951
VERSION AC130951.3 GI:25007317
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 259367)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D:, Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 259367)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (15-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 259367)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Nov 15, 2002 this sequence version replaced gi:23267482.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature

```
          table.
          -------------- Genome Center
              Center: Baylor College of Medicine
              Center code: BCM
              Web site: http://www.hgsc.bcm.tmc.edu/
              Contact: hgsc-help@bcm.tmc.edu
          -------------- Project Information
              Center project name: KCBZ
              Center clone name: CH230-28H16
          -------------- Summary Statistics
              Assembly program: Phrap; version 0.990329
              Consensus quality: 224177 bases at least Q40
              Consensus quality: 227857 bases at least Q30
              Consensus quality: 230151 bases at least Q20
              Estimated insert size: 228579; sum-of-contigs estimation
              Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
          --------------
          * NOTE: Estimated insert size may differ from sequence length
          *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 7 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *        1   223575: contig of 223575 bp in length
          *   223576   223675: gap of unknown length
          *   223676   251764: contig of 28089 bp in length
          *   251765   251864: gap of unknown length
          *   251865   252974: contig of 1110 bp in length
          *   252975   253074: gap of unknown length
          *   253075   254115: contig of 1041 bp in length
          *   254116   254215: gap of unknown length
          *   254216   255456: contig of 1241 bp in length
          *   255457   255556: gap of unknown length
          *   255557   256741: contig of 1185 bp in length
          *   256742   256841: gap of unknown length
          *   256842   259367: contig of 2526 bp in length.
FEATURES             Location/Qualifiers
     source          1..259367
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-28H16"
     misc_feature    928..6316
                     /note="clone_boundary
                     clone_end:T7
                     site:
                     end_sequence:BH365453"
     misc_feature    41559..43539
                     /note="wgs_contig"
     misc_feature    60167..61370
                     /note="wgs_contig"
     misc_feature    63848..65080
                     /note="wgs_contig"
     misc_feature    65401..67219
                     /note="wgs_contig"
ORIGIN

  Query Match             5.2%;  Score 159.6;  DB 2;  Length 259367;
  Best Local Similarity   54.9%;  Pred. No. 1.7e-18;
  Matches  704;  Conservative    0;  Mismatches  484;  Indels   94;  Gaps   16;

Qy        475 AAGAATAACAAATGCTATACCTCAGGAATATTAATTCCAGATTTACAGCATTTTAACTT 534
Db     123342 AAGACAAATCCAAGTTATTAAGATATAAACATAAACTCTGCACTCAAAGCCTTCTAATTT 123401

Qy        535 TCTTGATGAGAAAAAATAAATTTGTCAGTTATTAAACTATTTGGATCCAACAGATGAAAG 594
Db     123402 TATTGAAAAGGAAAAGTAAGCTTGTTAGCTATTAAACCAATGAGGATCCAGTACACAAAG 123461

Qy        595 CAGAATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGT 654
Db     123462 GACAATT---ACTGACTGACACGTTTCCTGATTTGTGACTTACGTATTTACATATGCTAA 123518

Qy        655 TTGACACAATT-----CTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTT 709
Db     123519 GTCTGACACATTAGCATTCACATTGGTTCTTGAGGAATACATATTTGCTTTTTGAATGTA 123578

Qy        710 AAGTTTCCTTTATTTTACTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAG 769
Db     123579 CA---TGTGTTTCTTCTTTTACTCTGTAGTCTTGGCCACATTTTTAAAGTTTTCTCCTTG 123635

Qy        770 ATAATTCAGCCTTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTTA 829
Db     123636 GAAATTTTAGATGTTAAGCAT----CCATTATAATCTCTG-CAGTTGGAACTATATTTTA 123690

Qy        830 ACTGTAAATATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCAC 889
Db     123691 A---TAAACATTTTTAACTGTTTCGTGATGGATTTCACTAGAACAACTTACATGCTCACA 123747

Qy        890 AAATCTTAAAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTT 949
Db     123748 TGTTTTAAAAGGCGTTCCTCAGAGCCTGTCTTGCTGTGCACAGTTGCTGCTTTGTTTGGT 123807

Qy        950 CTTTGCATGCCATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTGAATGATTGAG 1009
Db     123808 GTTATGAATCCGCACACATGACAGCTTCAGGGCTTCCCTATCCTCAAATGATCGAGCC-- 123865

Qy       1010 TGGGACTGTGCCAGCTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACACAC 1069
Db     123866 -GGGACTGTGCC----------------------CACGTCAGTCCATCTTCAGGTACAC 123901

Qy       1070 TGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTGCTC 1129
Db     123902 AGCCCTGGGTTAT-ATATTTCTGCCCAACTGGACTTTCTTTGGTGCTTTAGTGATTACTC 123960

Qy       1130 TTTTGTTTGGGCTTGGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACACAAT 1189
Db     123961 TTTTCTTT-GGTTTTAAGGTTTCAAAACTATGACATCCAAAGCTCAGATGTCAACACAAC 124019

Qy       1190 ATTAAGATTATTGCAATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTTTTG 1249
Db     124020 AGTAAGACAATTGCAGTTTGTAGCGGGTCCTTGCCTGTTGTCTAGAATTGGCTGCTCTGA 124079

Qy       1250 AGATTGCATT-----CCTTTCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTG 1304
Db     124080 GAATTGCATTATGACTTCTTCAGTTCAGGTATGTGATCAGATTCTCTAAGCAGATCTTTT 124139

Qy       1305 TGTG----GTATGTTTACTGTAGTTGCTTAGAAGTCCATTCTTTACCAATGCTCAAATGT 1360
Db     124140 TGTGTGCTGTTTTTTTATTCTGATTAATTAGAAGTGCAATTCTT---TATCCTCAGTTGG 124196

Qy       1361 GATTAAATTTGTTTTCTTGTTAAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTT 1420
Db     124197 GATTGCTTTTGTTTTTCCAAAAATACAATCAGCTTAAAACAAATTCCTAAATGTCTCTAA 124256

Qy       1421 ATTTCAGTGATTTAACATTTCCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTT 1480
Db     124257 TTTTCAGTATTAAGTCATTTC------------CTTATGAAAAAGGAAATAATGCCCTCT 124304

Qy       1481 TCTCCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACC 1540
Db     124305 TCCCCCAATAAA-------ACATGTGAGTTCTATAGCACAAACTGTGACACCACAGCACA 124357

Qy       1541 AACACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTG 1600
Db     124358 AATAGCCGAAAATGTTGTTTGCCATGTTTCCTA---------------TAGAAGTTGTG 124401

Qy       1601 AATGACAGATACTTCAGGCTTTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAA 1660
Db     124402 AATGATAGATGTCATAAGCTTTCAAAGGGAGCTAGAACATACGCCAGGCTTGGTACAAAC 124461

Qy       1661 GGTTCAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAA 1720
```

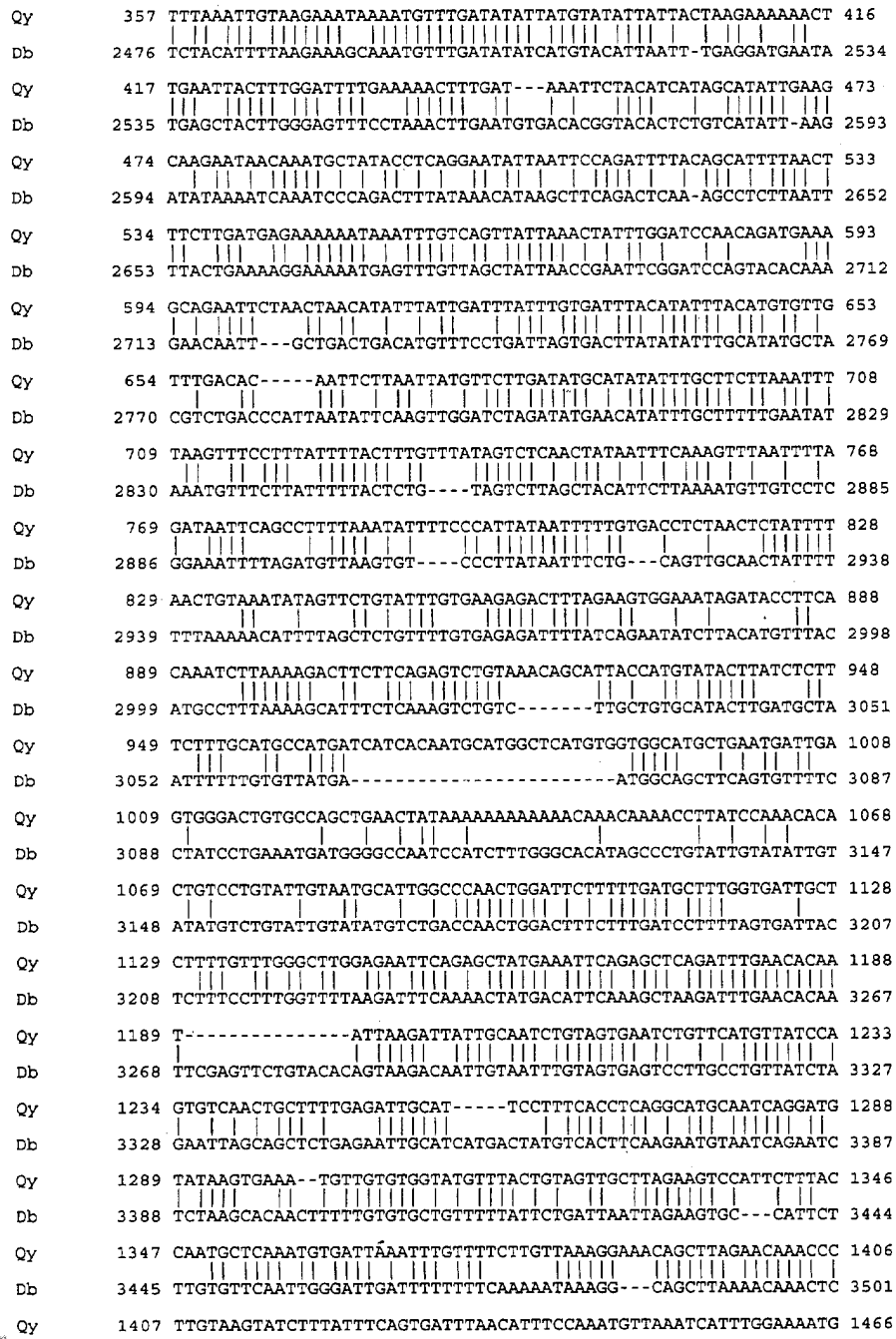
```
                   ||| || | | |  |  ||| || | |  | ||| |      || | |
Db     124462 ACCCAAGAAAAAGAAAAGAAAGAGAAAAGAAATGATTGAAGAAAAAATTTAAAGTGACAT 124521

Qy       1721 CTCAATACTACCTTGTCTTTTA 1742
                | |    | |  | | |||||
Db     124522 TTAATGGATTCTATTTATTTTA 124543


RESULT 11
AC094978
LOCUS       AC094978               260103 bp    DNA     linear   HTG 09-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-6I9, *** SEQUENCING IN PROGRESS ***,
            4 unordered pieces.
ACCESSION   AC094978
VERSION     AC094978.5  GI:30467213
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 260103)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 260103)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 260103)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 9, 2003 this sequence version replaced gi:23101287.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GBVQ
                Center clone name: CH230-6I9
            -------------- Summary Statistics
                Assembly program: Atlas;
                Consensus quality: 229362 bases at least Q40
                Consensus quality: 233666 bases at least Q30
                Consensus quality: 236756 bases at least Q20
                Estimated insert size: 242426; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1   252640: contig of 252640 bp in length
            *   252641   252740: gap of unknown length
            *   252741   253907: contig of 1167 bp in length
            *   253908   254007: gap of unknown length
            *   254008   255647: contig of 1640 bp in length
            *   255648   255747: gap of unknown length
            *   255748   260103: contig of 4356 bp in length.
FEATURES             Location/Qualifiers
     source          1. .260103
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-6I9"
     misc_feature    1. .3393
                     /note="wgs_contig"
     misc_feature    248789. .250071
                     /note="wgs_contig"
     misc_feature    251148. .252640
                     /note="wgs_contig"
ORIGIN

  Query Match             5.2%;  Score 159.6;  DB 2;  Length 260103;
```

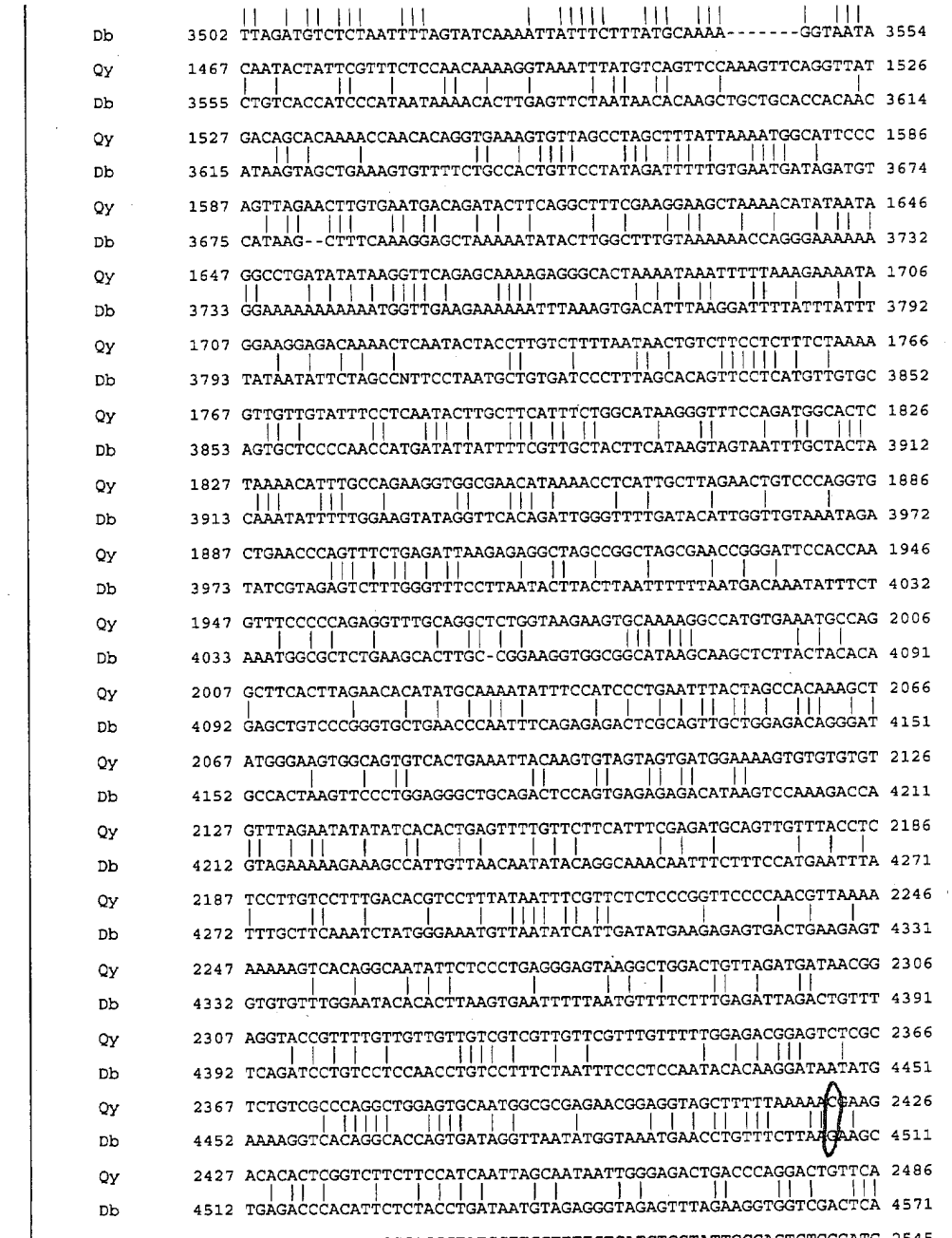
```
  Best Local Similarity   54.9%;  Pred. No. 1.7e-18;
  Matches  704;  Conservative    0;  Mismatches  484;  Indels   94;  Gaps   16;

Qy        475 AAGAATAACAAATGCTATACCTCAGGAATATTAATTCCAGATTTTACAGCATTTTAACTT 534
Db     112977 AAGACAAATCCAAGTTATTAAGATATAAACATAAACTCTGCACTCAAAGCCTTCTAATTT 113036

Qy        535 TCTTGATGAGAAAAAATAAATTTGTCAGTTATTAAACTATTTGGATCCAACAGATGAAAG 594
Db     113037 TATTGAAAAGGAAAAGTAAGCTTGTTAGCTATTAAACCAATGAGGATCCAGTACACAAAG 113096

Qy        595 CAGAATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGT 654
Db     113097 GACAATT---ACTGACTGACACGTTTCCTGATTTGTGACTTACGTATTTACATATGCTAA 113153

Qy        655 TTGACACAATT-----CTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTT 709
Db     113154 GTCTGACACATTAGCATTCACATTGGTTCTTGAGGAATACATATTTGCTTTTTGAATGTA 113213

Qy        710 AAGTTTCCTTTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAG 769
Db     113214 CA---TGTGTTTCTTCTTTTACTCTGTAGTCTTGGCCACATTTTTAAAGTTTTCTCCTTG 113270

Qy        770 ATAATTCAGCCTTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTTA 829
Db     113271 GAAATTTTAGATGTTAAGCAT----CCATTATAATCTCTG-CAGTTGGAACTATATTTTA 113325

Qy        830 ACTGTAAATATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCAC 889
Db     113326 A---TAAACATTTTTAACTGTTTCGTGATGGATTTCACTAGAACAACTTACATGCTCACA 113382

Qy        890 AAATCTTAAAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTT 949
Db     113383 TGTTTTAAAAGGCGTTCCTCAGAGCCTGTCTTGCTGTGCACAGTTGCTGCTTTGTTTGGT 113442

Qy        950 CTTTGCATGCCATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTGAATGATTGAG 1009
Db     113443 GTTATGAATCCGCACACATGACAGCTTCAGGGCTTCCCTATCCTCAAATGATCGAGCC-- 113500

Qy       1010 TGGGACTGTGCCAGCTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACACAC 1069
Db     113501 -GGGACTGTGCC----------CACGTCAGTCCATCTTCAGGTACAC 113536

Qy       1070 TGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTGCTC 1129
Db     113537 AGCCCTGGGTTAT-ATATTTCTGCCCAACTGGACTTTCTTTGGTGCTTTAGTGATTACTC 113595

Qy       1130 TTTTGTTTGGGCTTGGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACACAAT 1189
Db     113596 TTTTCTTT-GGTTTTAAGGTTTCAAAACTATGACATCCAAAGCTCAGATGTCAACACAAC 113654

Qy       1190 ATTAAGATTATTGCAATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTTTTG 1249
Db     113655 AGTAAGACAATTGCAGTTTGTAGCGGGTCCTTGCCTGTTGTCTAGAATTGGCTGCTCTGA 113714

Qy       1250 AGATTGCATT-----CCTTTCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTG 1304
Db     113715 GAATTGCATTATGACTTCTTCAGTTCAGGTATGTGATCAGATTCTCTAAGCAGATCTTTT 113774

Qy       1305 TGTG----GTATGTTTACTGTAGTTGCTTAGAAGTCCATTCTTTACCAATGCTCAAATGT 1360
Db     113775 TGTGTGCTGTTTTTTTATTCTGATTAATTAGAAGTGCAATTCTT---TATCCTCAGTTGG 113831

Qy       1361 GATTAAATTTGTTTTCTTGTTAAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTT 1420
Db     113832 GATTGCTTTTGTTTTTCCAAAAATACAATCAGCTTAAAACAAATTCCTAAATGTCTCTAA 113891

Qy       1421 ATTTCAGTGATTTAACATTTCCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTT 1480
Db     113892 TTTTCAGTATTAAGTCATTTC----------CTTATGAAAAAGGAAATAATGCCCTCT 113939

Qy       1481 TCTCCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACC 1540
```

```
Db     113940 TCCCCCAATAAA-------ACATGTGAGTTCTATAGCACAAACTGTGACACCACAGCACA 113992

Qy       1541 AACACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTG 1600
Db     113993 AATAGCCGAAAATGTTGTTTGCCATGTTTCCTA----------------TAGAAGTTGTG 114036

Qy       1601 AATGACAGATACTTCAGGCTTTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAA 1660
Db     114037 AATGATAGATGTCATAAGCTTTCAAAGGGAGCTAGAACATACGCCAGGCTTGGTACAAAC 114096

Qy       1661 GGTTCAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAA 1720
Db     114097 ACCCAAGAAAAAGAAAAGAAAGAGAAAAGAAATGATTGAAGAAAAAATTTAAAGTGACAT 114156

Qy       1721 CTCAATACTACCTTGTCTTTTA 1742
Db     114157 TTAATGGATTCTATTTATTTTA 114178
```

RESULT 12
AY029704

```
LOCUS       AY029704                5570 bp    DNA     linear   ROD 01-SEP-2001
DEFINITION  Mus musculus serotonin 1A receptor gene, promoter region and
            partial cds.
ACCESSION   AY029704
VERSION     AY029704.1  GI:15418779
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 5273 to 5560)
  AUTHORS   Parks,C.L. and Shenk,T.
  TITLE     The serotonin 1a receptor gene contains a TATA-less promoter that
            responds to MAZ and Sp1
  JOURNAL   J. Biol. Chem. 271 (8), 4417-4430 (1996)
  MEDLINE   96224025
   PUBMED   8626793
REFERENCE   2  (bases 1 to 5570)
  AUTHORS   Ansorge,M., Tanneberger,C., Davies,B., Theuring,F. and Kusserow,H.
  TITLE     Functional analysis of the murine 5-HT1A receptor promoter in vitro
            and in vivo
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 5570)
  AUTHORS   Kusserow,H., Ansorge,M.S. and Theuring,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2001) Institute for Pharmacology and Toxicology,
            Charite University Hospital, Dorotheenstr. 94, Berlin 10117,
            Germany
FEATURES             Location/Qualifiers
     source          1. .5570
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129/Sv"
                     /db_xref="taxon:10090"
                     /chromosome="13"
     promoter        1. .5272
     mRNA            5273. .>5570
                     /product="serotonin 1A receptor"
                     /citation=[1]
     CDS             5561. .>5570
                     /note="G protein coupled-receptor; 5-HT 1A receptor"
                     /codon_start=1
                     /product="serotonin 1A receptor"
                     /protein_id="AAK50428.1"
                     /db_xref="GI:15418780"
                     /translation="MDM"
ORIGIN

  Query Match             4.7%;  Score 144;  DB 10;  Length 5570;
  Best Local Similarity   48.0%;  Pred. No. 1.8e-15;
  Matches 1326;  Conservative    0;  Mismatches 1286;  Indels  151;  Gaps   26;
```





```
Db       4572 GCTTCCCATCCTTCAGAGTCTCTGAGGGTTTTCCTCGTGCCTGATAGCCACACTGTGAAC 4631

Qy       2546 CTGACACGATTTAAGAATTTGGCAGATAATATGAGGCAAGGAG----------------- 2588

Db       4632 CTGGCTCTATTTAAGAAGCTGACAGAAATAAAGGGGATCTGCGGTGGGGAAATATATCTT 4691

Qy       2589 --------------TAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTTTTGCT 2633

Db       4692 TCTATCTCTTTCTTTTATTTTATTTTTCCTTTTGTCTTTTTTCCCCCTCTAGTTTAATA 4751

Qy       2634 GGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGATCTGGTGTGTG 2693

Db       4752 GGTAAGAGGCGGGGTTTAGTATGTTAAAACATTGATCTGACTAGCAGATCAGAGTAGATA 4811

Qy       2694 T---AAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTGAGC 2750

Db       4812 TATGAAGTGAATTCTGAGTGTGGTGTG---------GACTGGAATGTAAAGACTCCTGGC 4862

Qy       2751 TAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAAGTA 2810

Db       4863 TACAGGGAGAGGAGGGCGGGGACCCAGAAGAAGGAGGCACTCGGGACTAGGGAAGTGTTG 4922

Qy       2811 TTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGACTCACAGAG 2870

Db       4923 GATGGCGGTGGGGGTTAACGTGGGAGGAAGGGAGACTAGCTGTCTGAGCGACACAC--AG 4980

Qy       2871 GGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGGTGGGGAGA 2930

Db       4981 GACTAAAGAAAGAGAGGTGAGAAGGAAAGGGGAACTGGAA--GGAAGGCCAGTGGGAGGA 5038

Qy       2931 AGGGGGACGAAAGAGGCAGAAGAGAGAGAA--------GAGAGGAGGAGAGAGGGGGAGA 2982

Db       5039 AGCCGGGAGAAGGAAGCTGGGAAAAGAGAAGCAGACTGGAGAAGGCAGGAGAGATGGAGA 5098

Qy       2983 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 3042

Db       5099 GACAGAGACGAAGGAACTAGGAGATGAGGGTCATCGAGTGACCCGGGAGGATGGGGTTTC 5158

Qy       3043 TCG 3045

Db       5159 TGG 5161
```

RESULT 13
AF217200
LOCUS AF217200 2725 bp DNA linear ROD 13-MAR-2000
DEFINITION Rattus norvegicus serotonin 5-HT1A receptor gene, partial cds.
ACCESSION AF217200
VERSION AF217200.1 GI:6969399
KEYWORDS .
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 2725)
AUTHORS Storring,J.M., Charest,A., Cheng,P. and Albert,P.R.
TITLE TATA-driven transcriptional initiation and regulation of the rat serotonin 5-HT1A receptor gene
JOURNAL J. Neurochem. 72 (6), 2238-2247 (1999)
MEDLINE 99277334
PUBMED 10349831
REFERENCE 2 (bases 1 to 2725)
AUTHORS Ou,X.M., Jafar-Nejad,H., Storring,J.M., Meng,J.H., Lemonde,S. and Albert,P.R.
TITLE Novel dual repressor elements for neuronal cell-specific transcription of the rat 5-HT1A receptor gene
JOURNAL J. Biol. Chem. 275 (11), 8161-8168 (2000)
MEDLINE 20179926
PUBMED 10713139
REFERENCE 3 (bases 1 to 2725)
AUTHORS Storring,J.M. and Albert,P.R.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1999) Neuroscience Res. Inst., University of Ottawa, 451 Smyth Road, Ottawa, ON K1H-8M5, Canada
FEATURES Location/Qualifiers
source 1. .2725
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
repeat_region 1447. .1499
/rpt_type=tandem
/rpt_unit="gt"
mRNA 1763. .>2725
/product="serotonin 5-HT1A receptor"
CDS 2720. .>2725
/note="G protein-coupled receptor"
/codon_start=1
/product="serotonin 5-HT1A receptor"
/protein_id="AAF33756.1"
/db_xref="GI:6969400"
/translation="MD"
ORIGIN

Query Match 4.6%; Score 139.6; DB 10; Length 2725;
Best Local Similarity 57.9%; Pred. No. 1.2e-14;
Matches 402; Conservative 0; Mismatches 249; Indels 43; Gaps 7;

```
Qy       1054 CCTTATCCAAACACACTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGAT 1113

Db        218 CCATCTTCAGGTACACAGCCCTGGGTTAT-ATATTTCTGCCCAACTGGACTTTCTTTGGT 276

Qy       1114 GCTTTGGTGATTGCTCTTTTGTTTGGGCTTGGAGAATTCAGAGCTATGAAATTCAGAGCT 1173

Db        277 GCTTTAGTGATTACTCTTTTCTTTGGTTTTA-AGGTTTCAAAACTATGACATCCAAAGCT 335

Qy       1174 CAGATTTGAACACAATATTAAGATTATTGCAATCTGTAGTGAATCTGTTCATGTTATCCA 1233

Db        336 CAGATGTCAACACAACAGTAAGACAATTGCAGTTTGTAGCGGGTCCTTGCCTGTTGTCTA 395

Qy       1234 GTGTCAACTGCTTTTGAGATTGCATT-----CCTTTCACCTCAGGCATGCAATCAGGATG 1288

Db        396 GAATTGGCTGCTCTGAGAATTGCATTATGACTTCTTCAGTTCAGGTATGTGATCAGATTC 455

Qy       1289 TATAAGTGAAATGTTGTGTGGTATGTTTACTGTAGTTGCTTAGAAGTCCATTCTTTACCA 1348

Db        456 TCTAAG-CAGATCTTTTGTGTGCTGTTTTTTATTCTGATTAATTAGAAGTGCAATTCTTT 514

Qy       1349 ATGCTCAAATGTGATTAAATTTGTTTTCTTGTTAAAGGAAACAGCTTAGAACAAACCCTT 1408

Db        515 ATCCTCAGTTGGGATTGCTTTTGTTTTTCCAAAAATACAATCAGCTTAAAACAAATTCCT 574

Qy       1409 GTAAGTATCTTTATTTCAGTGATTTAACATTTCCAAATGTTAAATCATTTGGAAAATGCA 1468

Db        575 AAATGTCTCTAATTTTCAGTATTAAGTCATTTC------------CTTATGAAAAAGGAA 622

Qy       1469 ATACTATTCGTTTCTCCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTCAGGTTATGA 1528

Db        623 ATAATGCCCTCTTCCCCCAATAAA-------ACATGTGAGTTCTATAGCACAAACTGTGA 675

Qy       1529 CAGCACAAAACCAACACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGGCATTCCCAG 1588

Db        676 CACCACAGCACAAATAGCCGAAAATGTTGTTTGCCATGTTTCCTA--------------- 720

Qy       1589 TTAGAACTTGTGAATGACAGATACTTCAGGCTTTCGAAGGAAGCTAAAACATATAATAGG 1648

Db        721 -TAGAAGTTGTGAATGATAGATGTCATAAGCTTTCAAAGGGAGCTAGAACATACGCCAGG 779

Qy       1649 CCTGATATATAAGGTTCAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATAGG 1708

Db        780 CTTGGTACAAACACCCAAGAAAAAGAAAAGAAAGAGAAAAGAAATGATTGAAGAAAAAAT 839

Qy       1709 AAGGAGACAAAACTCAATACTACCTTGTCTTTTA 1742

Db        840 TTAAAGTGACATTTAATGGATTCTATTTATTTTA 873
```

```
RESULT 14
MMU33820
LOCUS       MMU33820                1256 bp    DNA     linear   ROD 30-MAY-1996
DEFINITION  Mus musculus serotonin 1a receptor (Htr1a) gene, 5' flanking
            sequence and partial cds.
ACCESSION   U33820
VERSION     U33820.1  GI:1335803
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1256)
  AUTHORS   Parks,C.L. and Shenk,T.
  TITLE     The serotonin 1a receptor gene contains a TATA-less promoter that
            responds to MAZ and Sp1
  JOURNAL   J. Biol. Chem. 271 (8), 4417-4430 (1996)
  MEDLINE   96224025
   PUBMED   8626793
REFERENCE   2  (bases 1 to 1256)
  AUTHORS   Parks,C.L. and Shenk,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-AUG-1995) Christopher L. Parks, Molecular Biology,
            Princeton University, Washington Road, Princeton, NJ 08544-1014,
            USA
FEATURES             Location/Qualifiers
     source          1. .1256
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C3H/An"
                     /db_xref="taxon:10090"
                     /cell_type="L-cell"
     misc_feature    511. .1175
                     /note="multiple mRNA 5' ends"
     gene            1188. .1256
                     /gene="Htr1a"
     CDS             1188. .>1256
                     /gene="Htr1a"
                     /note="5-HT1a receptor; neurotransmitter receptor"
                     /codon_start=1
                     /product="serotonin 1a receptor"
                     /protein_id="AAC52572.1"
                     /db_xref="GI:1335804"
                     /translation="MDMFSLGQGNNTTTSLDHLGTGG"
ORIGIN

  Query Match             4.4%;  Score 134.2;  DB 10;  Length 1256;
  Best Local Similarity   64.7%;  Pred. No. 1.2e-13;
  Matches  288;  Conservative    0;  Mismatches  133;  Indels   24;  Gaps    5;

Qy       2612 TTTTTCCAACCCCAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCT 2671
             ||||||| | | |||||    ||   |  |  ||  ||  | ||||| ||| ||| |||
Db        357 TTTTTCCCCCTCTAGTTTAATAGGTAAGAGGCGGGGTTTAGTATGTTAAAACATTGATCT 416

Qy       2672 GACCGGCAGGATCTGGTGTGTGT---AAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGA 2728
             |||  ||||       ||   | |   |||||| |||||||| |    ||          ||
Db        417 GACTAGCAGATCAGAGTAGATATATGAAGTGAATTCTGAGTGTGGTGTG---------GA 467

Qy       2729 CTCGAATGCAAAGACGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGC 2788
             || ||||| |||||| |    |||| ||||||||||||||||||||||||| ||| |||||
Db        468 CTGGAATGTAAAGACTCCTGGCTACAGGGAGAGGAGGGCGGGGACCCAGAAGAAGGAGGC 527

Qy       2789 ACTCCTCGGGGTTGGGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTG 2848
             ||||          |  |   |    || || | ||||||  |||||||| | |||| ||
Db        528 ACTCGGGACTAGGGAAGTGTTGGATGGCGGTGGGGGTTAACGTGGGAGGAAGGGAGACTA 587

Qy       2849 GCTTTCGAAGCGACTCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTG 2908
             ||| ||  |||||  |||| |||  |||| |||| || ||||| |||||  ||||    |
Db        588 GCTGTCTGAGCGA--CACACAGGACTAAAGAAAGAGAGGTGAGAAGGAAAGGGGAACTGG 645

Qy       2909 AAAGGGAAGGCAGGTGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAA-------- 2960
                ||  |||||||  |||||  ||||  ||  |||  || || |   | ||||||
```

```
Db        646 AA--GGAAGGCCAGTGGGAGGAAGCCGGGAGAAGGAAGCTGGGAAAAGAGAAGCAGACTG 703

Qy       2961 GAGAGGAGGAGAGAGGGGGAGAGAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAG 3020
              |||| |    |||||  |||||||  ||    | | ||  |||||  ||||||||  |||
Db        704 GAGAAGGCAGGAGAGATGGAGAGACAGAGACGAAGGAACTAGGAGATGAGGGTCATCGAG 763

Qy       3021 TGACCGTGGAGGATGGGGCTTCTCG 3045
              |||||  |||||||||||| |||| |
Db        764 TGACCCGGGAGGATGGGGTTTCTGG 788


RESULT 15
I66494/c
LOCUS       I66494                  7218 bp    DNA     linear   PAT 28-DEC-1997
DEFINITION  Sequence 14 from patent US 5670367.
ACCESSION   I66494
VERSION     I66494.1  GI:2724471
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 7218)
  AUTHORS   Dorner,F., Scheiflinger,F. and Falkner,F.Gunter.
  TITLE     Recombinant fowlpox virus
  JOURNAL   Patent: US 5670367-A 14 23-SEP-1997;
FEATURES             Location/Qualifiers
     source          1. .7218
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             4.1%;  Score 125.6;  DB 6;  Length 7218;
  Best Local Similarity   2.1%;  Pred. No. 2.8e-12;
  Matches    8;  Conservative  284;  Mismatches   88;  Indels    0;  Gaps    0;

Qy       2665 TTGGTCTGACCGGCAGGATCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAA 2724
              |||||  ::  :: ::::   :: : : : ::: :::    :::     :  :: ::::::
Db        1441 TTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1382

Qy       2725 GAGACTCGAATGCAAAGACGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAG 2784
              ::::   ::: : ::::: :  :::  :::::::::::::::: :::::   :::::::::
Db        1381 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1322

Qy       2785 AGGCACTCCTCGGGGTTGGGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG 2844
              ::: :      ::::  ::::::: :  :::::::::::  :::: ::::::::::::::
Db        1321 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1262

Qy       2845 CCTGGCTTTCGAAGCGACTCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAG 2904
                ::     :::: ::   : ::::::: ::: :::::::: ::::::::::::::::
Db        1261 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1202

Qy       2905 ACTGAAAGGGAAGGCAGGTGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGA 2964
              :  :::::::::: ::: :::::::::::::: :::::::: ::::::::::::::::::
Db        1201 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1142

Qy       2965 GGAGGAGAGAGGGGGAGAGAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGAC 3024
              :::::::::::::::::::::::::::::::::::: :::::::::::::  : :::: ::
Db        1141 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1082

Qy       3025 CGTGGAGGATGGGGCTTCTC 3044
              : :::::: ::::  || ||
Db        1081 RRRRRRRRRRRRRRRATCGC 1062


Search completed: April 10, 2004, 12:31:28
Job time : 11446 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 04:24:23 ; Search time 1088 Seconds
(without alignments)
11889.488 Million cell updates/sec

Title: US-09-430-412A-1
Perfect score: 3045
Sequence: 1 atcatcaataatatccgtta.....gtggaggatggggcttctcg 3045

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	1959.2	64.3	3000	4	AAH88467	Aah88467 CNS disor
	2	979.4	32.2	3001	4	AAH88468	Aah88468 CNS disor
	3	747.8	24.6	1204	5	AAF75001	Aaf75001 Part of H
	4	747.8	24.6	2722	5	AAF75003	Aaf75003 Part of H
	5	742.4	24.4	839	3	AAC71827	Aac71827 Single nu
	6	742.4	24.4	839	3	AAC71824	Aac71824 Single nu
	7	742.4	24.4	839	3	AAC71830	Aac71830 Single nu
	8	522.4	17.2	2512	4	AAH88477	Aah88477 CNS disor
	9	89.2	2.9	8056	7	ABZ10246	Abz10246 Haematopo
	10	86.8	2.9	8056	7	ABZ10100	Abz10100 Haematopo
	11	83.2	2.7	11745	6	ABK28332	Abk28332 DNA trans
	12	82.4	2.7	6636	6	ABL32791	Abl32791 Human imm
c	13	81.4	2.7	105325	6	ABK94407	Abk94407 DNA encod
c	14	81.2	2.7	8056	7	ABZ10246	Abz10246 Haematopo
c	15	80.6	2.6	700	4	AAH93026	Aah93026 Human inf
	16	78.8	2.6	19787	6	ABL33451	Abl33451 Human imm
	17	77.8	2.6	6063	6	ABK28394	Abk28394 DNA trans
	18	77	2.5	19634	7	ABZ10161	Abz10161 Haematopo
c	19	76.2	2.5	8056	7	ABZ10100	Abz10100 Haematopo
	20	75.2	2.5	6317	6	ABL32408	Abl32408 Human imm
	21	75.2	2.5	6317	6	ABL49311	Abl49311 Human pol
	22	75.2	2.5	6465	6	ABL32985	Abl32985 Human imm
	23	75.2	2.5	7351	6	ABL32029	Abl32029 Human imm
	24	75.2	2.5	8222	7	ACF62816	Acf62816 Colon can
	25	75.2	2.5	11222	9	ADB54318	Adb54318 Pretreate
	26	74.8	2.5	5145	6	ABL32348	Abl32348 Human imm
	27	74.8	2.5	5145	6	ABL34464	Abl34464 Human met
	28	74.4	2.4	14919	4	AAS46506	Aas46506 Tumour su
c	29	74	2.4	665	3	AAZ31996	Aaz31996 Human apo
	30	74	2.4	5689	4	AAS45384	Aas45384 Chemicall
	31	74	2.4	5689	4	AAS46426	Aas46426 Tumour su
	32	74	2.4	5689	6	ABK28226	Abk28226 DNA trans
	33	74	2.4	12237	6	ABL34358	Abl34358 Human imm
	34	73.8	2.4	19634	7	ABZ10015	Abz10015 Haematopo
	35	73.6	2.4	4985	6	ABQ75107	Abq75107 Anopheles
	36	73.6	2.4	4985	9	ACF79720	Acf79720 Mosquito
	37	73.6	2.4	8222	7	ACF62794	Acf62794 Colon can
	38	73.6	2.4	11222	9	ADB54190	Adb54190 Pretreate
	39	73.4	2.4	8170	6	ABK28258	Abk28258 DNA trans
	40	73.2	2.4	18585	6	ABL34609	Abl34609 Human met
	41	72.6	2.4	115218	7	ACA64845	Aca64845 Human HNR
	42	72.2	2.4	6775	6	ABQ67160	Abq67160 Human ang
	43	72.2	2.4	32392	6	ABL56203	Abl56203 AmEPV gen
c	44	72	2.4	2503	2	AAQ53480	Aaq53480 pNPX30 xy
	45	72	2.4	5511	6	ABL33871	Abl33871 Human imm

ALIGNMENTS

RESULT 1
AAH88467/c
ID AAH88467 standard; DNA; 3000 BP.
XX
AC AAH88467;
XX
DT 26-FEB-2002 (first entry)
XX
DE CNS disorder-related biallelic marker #1 from 5HT1A gene.
XX
KW Single nucleotide polymorphism; SNP; biallelic marker; human;
KW central nervous system disorder; CNS; serotonin receptor; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_feature 1501
FT /*tag= a
FT /standard_name= "single nucleotide polymorphism"
XX
PN WO200151659-A2.
XX
PD 19-JUL-2001.
XX
PF 11-JAN-2001; 2001WO-IB000116.
XX
PR 13-JAN-2000; 2000US-0175854P.
XX
PA (GEST) GENSET.
XX
PI Chu T, Blumenfeld M, Cohen D;
XX
DR WPI; 2001-483085/52.
XX
PT Isolated polynucleotides, useful for genotyping nucleic acids for
PT biallelic markers for the diagnosis of depression, comprises central
PT nervous system disorder related biallelic marker.
XX
PS Claim 1; Page 335-336; 519pp; English.
XX
CC The present invention relates to biallelic markers derived from human
CC genes involved in central nervous system (CNS) disorders. The present
CC sequence is one such biallelic marker derived from a human serotonin
CC receptor gene. This marker has a single nucleotide polymorphism (SNP) and
CC is useful in determining the genetic predisposition of individuals to CNS
CC disorders, by identifying the nucleotides at a set of genetic markers in

CC a biological sample, where the markers comprise at least one CNS disorder
CC related marker
XX
SQ Sequence 3000 BP; 783 A; 780 C; 665 G; 771 T; 0 U; 1 Other;

Query Match 64.3%; Score 1959.2; DB 4; Length 3000;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2018; Conservative 1; Mismatches 1; Indels 6; Gaps 6;

```
Qy    1024 CTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACACACTGTCCTGTATTGTA 1083
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2900 CTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACACACTGTCCTGTATTGTA 2841

Qy    1084 ATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTGCTCTTTTGTTTGGGCTT 1143
           |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    2840 ATGCATTGGCCCAAGTGGATTCTTTTTGATGCTTTGGTGATTGCTCTTTTGTTTGGGCTT 2781

Qy    1144 GGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACACAATATTAAGATTATTGC 1203
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2780 GGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACACAATATTAAGATTATTGC 2721

Qy    1204 AATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTTTTGAGATTGCATTCCTT 1263
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2720 AATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTTTTGAGATTGCATTCCTT 2661

Qy    1264 TCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTGTGTGGTATGTTTACTGTAG 1323
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2660 TCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTGTGTGGTATGTTTACTGTAG 2601

Qy    1324 TTGCTTAGAAGT-CCATTCTTTACCAATGCTCAAATGTGATTAAATTTG-TTTTCTTGTT 1381
           |||||||||||| ||||||||||||||||||||||||||||||||| ||||||||||
Db    2600 TTGCTTAGAAGTCCCATTCTTTACCAATGCTCAAATGTGATTAAATTTGTTTTTCTTGTT 2541

Qy    1382 AAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTTA-TTTCAGTGATTTAACATTT 1440
           |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db    2540 AAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTTATTTTCAGTGATTTAACATTT 2481

Qy    1441 CCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTTTCTCCAACAAAAGGTAAATT 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2480 CCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTTTCTCCAACAAAAGGTAAATT 2421

Qy    1501 TATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACCAACACAGGTGAAAGTGTTAG 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2420 TATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACCAACACAGGTGAAAGTGTTAG 2361

Qy    1561 CCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTGAATGACAGATACTTCAGGCT 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2360 CCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTGAATGACAGATACTTCAGGCT 2301

Qy    1621 TTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAAGGTTCAGAGCAAAAGAGGGC 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2300 TTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAAGGTTCAGAGCAAAAGAGGGC 2241

Qy    1681 ACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAACTCAATACTACCTTGTCTTT 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2240 ACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAACTCAATACTACCTTGTCTTT 2181

Qy    1741 TAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCAATACTTGCTTCATTTCT 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2180 TAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCAATACTTGCTTCATTTCT 2121

Qy    1801 GGCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAAGGTGGCGAACATAAAAC 1860
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    2120 -GCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAAGGTGGCG-ACATAAAAC 2063

Qy    1861 CTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGAGATTAAGAGAGGCTAGC 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2062 CTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGAGATTAAGAGAGGCTAGC 2003

Qy    1921 CGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTTGCAGGCTCTGGTAAGAA 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2002 CGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTTGCAGGCTCTGGTAAGAA 1943
```

```
Qy    1981 GTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACATATGCAAAATATTTCCA 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1942 GTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACATATGCAAAATATTTCCA 1883

Qy    2041 TCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGTCACTGAAATTACAAGTG 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1882 TCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGTCACTGAAATTACAAGTG 1823

Qy    2101 TAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCACACTGAGTTTTGTTCTT 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1822 TAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCACACTGAGTTTTGTTCTT 1763

Qy    2161 CATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACACGTCCTTTATAATTTCGT 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1762 CATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACACGTCCTTTATAATTTCGT 1703

Qy    2221 TCTCTCCCGGTTCCCCAACGTTAAAAAAAAAAGTCACAGGCAATATTCTCCCTGAGGGAGT 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1702 TCTCTCCCGGTTCCCCAACGTTAAAAAAAAAAGTCACAGGCAATATTCTCCCTGAGGGAGT 1643

Qy    2281 AAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGTTGTTGTCGTCGTTGTTC 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1642 AAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGTTGTTGTCGTCGTTGTTC 1583

Qy    2341 GTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAATGGCGCGAGA 2400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1582 GTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAATGGCGCGAGA 1523

Qy    2401 ACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATCAATTAGCAATAA 2460
           ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db    1522 ACGGAGGTAGCTTTTTAAAAASGAAGACACACTCGGTCTTCTTCCATCAATTAGCAATAA 1463

Qy    2461 TTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTATGCTTCCTTTTCT 2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1462 TTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTATGCTTCCTTTTCT 1403

Qy    2521 CATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGCAGATAATATGAG 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1402 CATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGCAGATAATATGAG 1343

Qy    2581 GCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTTTTGCTGGGTTGG 2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1342 GCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTTTTGCTGGGTTGG 1283

Qy    2641 AGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGATCTGGTGTGTGTAAGTG 2699
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    1282 AGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGATCTGGTGTGTGTAAGTG 1223

Qy    2700 AGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTGAGCTAGAGGGAG 2759
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1222 AGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTGAGCTAGAGGGAG 1163

Qy    2760 AGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAAGTATTAGGAGGG 2819
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1162 AGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAAGTATTAGGAGGG 1103

Qy    2820 GAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGACTCACAGAGGGATAAATA 2879
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1102 GAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGACTCACAGAGGGATAAATA 1043

Qy    2880 AAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGGTGGGGAGAAGGGGGACG 2939
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1042 AAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGGTGGGGAGAAGGGGGACG 983

Qy    2940 AAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGGAAGGAAGGAAAT 2999
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     982 AAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGGAAGGAAGGAAAT 923

Qy    3000 AGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 3045
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     922 AGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 877
```

```
RESULT 2
AAH88468
ID   AAH88468 standard; DNA; 3001 BP.
XX
AC   AAH88468;
XX
DT   26-FEB-2002  (first entry)
XX
DE   CNS disorder-related biallelic marker #2 from 5HT1A gene.
XX
KW   Single nucleotide polymorphism; SNP; biallelic marker; human;
KW   central nervous system disorder; CNS; serotonin receptor; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1501
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200151659-A2.
XX
PD   19-JUL-2001.
XX
PF   11-JAN-2001; 2001WO-IB000116.
XX
PR   13-JAN-2000; 2000US-0175854P.
XX
PA   (GEST ) GENSET.
XX
PI   Chu T,  Blumenfeld M,  Cohen D;
XX
DR   WPI; 2001-483085/52.
XX
PT   Isolated polynucleotides, useful for genotyping nucleic acids for
PT   biallelic markers for the diagnosis of depression, comprises central
PT   nervous system disorder related biallelic marker.
XX
PS   Claim 1; Page 337; 519pp; English.
XX
CC   The present invention relates to biallelic markers derived from human
CC   genes involved in central nervous system (CNS) disorders. The present
CC   sequence is one such biallelic marker derived from a human serotonin
CC   receptor gene. This marker has a single nucleotide polymorphism (SNP) and
CC   is useful in determining the genetic predisposition of individuals to CNS
CC   disorders, by identifying the nucleotides at a set of genetic markers in
CC   a biological sample, where the markers comprise at least one CNS disorder
CC   related marker
XX
SQ   Sequence 3001 BP; 654 A; 799 C; 856 G; 678 T; 0 U; 14 Other;

  Query Match           32.2%;  Score 979.4;  DB 4;  Length 3001;
  Best Local Similarity   98.7%;  Pred. No. 7.2e-176;
  Matches 1005;  Conservative    0;  Mismatches   11;  Indels    2;  Gaps    2;

Qy       2029 AAAATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGTCACTG 2088
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 AANATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGTCACTG 68

Qy       2089 AAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCACACTG 2148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db         69 AAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCACANTG 128

Qy       2149 AGTTTTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACACGTCCT 2208
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 AGTTNTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACACGTCCT 188

Qy       2209 TTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAAGTCACAGGCAATATTCT 2268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 TTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAAGTCACAGGCAATATTCT 248

Qy       2269 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGTTGTTG 2328
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||

Db        249 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGNAGGTACCGTTTTGTTGTTGTTG 308

Qy       2329 TCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGC 2388
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 TCGTCGNTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGC 368

Qy       2389 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATC 2448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATC 428

Qy       2449 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 2508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 488

Qy       2509 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 2568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 548

Qy       2569 AGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 2628
              |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        549 AGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 608

Qy       2629 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGATCTGG 2687
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||||
Db        609 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGACCTGG 668

Qy       2688 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 2747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 728

Qy       2748 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 2807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 788

Qy       2808 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGACTCACA 2867
              |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||||
Db        789 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCGCTTTCGAAGCGACTCACA 847

Qy       2868 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGGTGGGG 2927
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        848 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTAAAGGGAAGGCAGGTGGGG 907

Qy       2928 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 2987
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        908 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 967

Qy       2988 AAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 3045
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        968 AAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 1025

RESULT 3
AAF75001
ID   AAF75001 standard; DNA; 1204 BP.
XX
AC   AAF75001;
XX
DT   08-MAY-2001  (first entry)
XX
DE   Part of HTR1A gene #1.
XX
KW   5-hydroxy tryptamine receptor 1A; HTR1A; polymorphism; Tourette's;
KW   neuropsychiatric; ds.
XX
OS   Homo sapiens.
XX
PN   WO200110884-A1.
XX
PD   15-FEB-2001.
XX
PF   01-AUG-2000; 2000WO-US040519.
XX
PR   06-AUG-1999;   99US-0147711P.
```

```
XX
PA   (GENA-) GENAISSANCE PHARM INC.
XX
PI   Denton RR,  Kliem SE,  Nandabalan K,  Stephens JC;
XX
DR   WPI; 2001-191514/19.
XX
PT   New 5-hydroxy tryptamine receptor 1A gene variants for studying
PT   expression and biological function of the gene and for developing drugs
PT   targeting 5-hydroxy tryptamine receptor 1A protein.
XX
PS   Disclosure; Fig 1; 64pp; English.
XX
CC   The present invention relates to 5-hydroxy tryptamine receptor 1A (HTR1A)
CC   gene. HTR1A-encoding polynucleotides containing one or more of the novel
CC   polymorphic sites are useful in studying the expression and biological
CC   function of HTR1A, as well as in developing drugs targeting this protein.
CC   In addition, information on the combinations of polymorphisms in the
CC   HTR1A gene may have diagnostic and forensic applications. A polymorphic
CC   variant of HTR1A is useful in studying the effect of the variation on the
CC   biological activity of HTR1A as well as studying the binding affinity of
CC   candidate drugs targeting HTR1A for the treatment of neuropsychiatric
CC   diseases and Tourette's syndrome
XX
SQ   Sequence 1204 BP; 283 A; 266 C; 394 G; 261 T; 0 U; 0 Other;

  Query Match             24.6%;  Score 747.8;  DB 5;  Length 1204;
  Best Local Similarity   98.9%;  Pred. No. 4.3e-132;
  Matches  774;  Conservative    0;  Mismatches    7;  Indels    2;  Gaps    2;

Qy       2264 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 2323
             ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAGCGGAGGTACCGTTTTGTTGT 60

Qy       2324 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 2383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 120

Qy       2384 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 2443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 180

Qy       2444 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 2503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 240

Qy       2504 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 2563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 300

Qy       2564 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 2623
             ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTGGCAGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 360

Qy       2624 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGA 2682
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        361 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGA 420

Qy       2683 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 480

Qy       2743 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 2802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 540

Qy       2803 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 2862
             |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||||
Db        541 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCGCTTTCGAAGCGAC 599

Qy       2863 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 2922
             ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        600 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTAAAGGGAAGGCAGG 659

Qy       2923 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 2982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 719

Qy       2983 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 3042
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        720 GAGGGAAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 779

Qy       3043 TCG 3045
             |||
Db        780 TCG 782


RESULT 4
AAF75003
ID   AAF75003 standard; DNA; 2722 BP.
XX
AC   AAF75003;
XX
DT   08-MAY-2001  (first entry)
XX
DE   Part of HTR1A gene #3.
XX
KW   5-hydroxy tryptamine receptor 1A; HTR1A; polymorphism; Tourette's;
KW   neuropsychiatric; ds.
XX
OS   Homo sapiens.
XX
PN   WO200110884-A1.
XX
PD   15-FEB-2001.
XX
PF   01-AUG-2000; 2000WO-US040519.
XX
PR   06-AUG-1999;   99US-0147711P.
XX
PA   (GENA-) GENAISSANCE PHARM INC.
XX
PI   Denton RR,  Kliem SE,  Nandabalan K,  Stephens JC;
XX
DR   WPI; 2001-191514/19.
XX
PT   New 5-hydroxy tryptamine receptor 1A gene variants for studying
PT   expression and biological function of the gene and for developing drugs
PT   targeting 5-hydroxy tryptamine receptor 1A protein.
XX
PS   Claim 1; Fig 3; 64pp; English.
XX
CC   The present invention relates to 5-hydroxy tryptamine receptor 1A (HTR1A)
CC   gene. HTR1A-encoding polynucleotides containing one or more of the novel
CC   polymorphic sites are useful in studying the expression and biological
CC   function of HTR1A, as well as in developing drugs targeting this protein.
CC   In addition, information on the combinations of polymorphisms in the
CC   HTR1A gene may have diagnostic and forensic applications. A polymorphic
CC   variant of HTR1A is useful in studying the effect of the variation on the
CC   biological activity of HTR1A as well as studying the binding affinity of
CC   candidate drugs targeting HTR1A for the treatment of neuropsychiatric
CC   diseases and Tourette's syndrome
XX
SQ   Sequence 2722 BP; 582 A; 743 C; 811 G; 586 T; 0 U; 0 Other;

  Query Match             24.6%;  Score 747.8;  DB 5;  Length 2722;
  Best Local Similarity   98.9%;  Pred. No. 5e-132;
  Matches  774;  Conservative    0;  Mismatches    7;  Indels    2;  Gaps    2;

Qy       2264 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 2323
             ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAGCGGAGGTACCGTTTTGTTGT 60

Qy       2324 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 2383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 120
```

```
Qy       2384 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 2443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 180

Qy       2444 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 2503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 240

Qy       2504 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 2563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 300

Qy       2564 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 2623
             |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTGGCAGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 360

Qy       2624 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGA 2682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        361 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGA 420

Qy       2683 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 480

Qy       2743 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 2802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 540

Qy       2803 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 2862
             |||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||
Db        541 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCGCTTTCGAAGCGAC 599

Qy       2863 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 2922
             ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        600 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTAAAGGGAAGGCAGG 659

Qy       2923 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 2982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 719

Qy       2983 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 3042
             |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        720 GAGGGAAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 779

Qy       3043 TCG 3045
             |||
Db        780 TCG 782
```

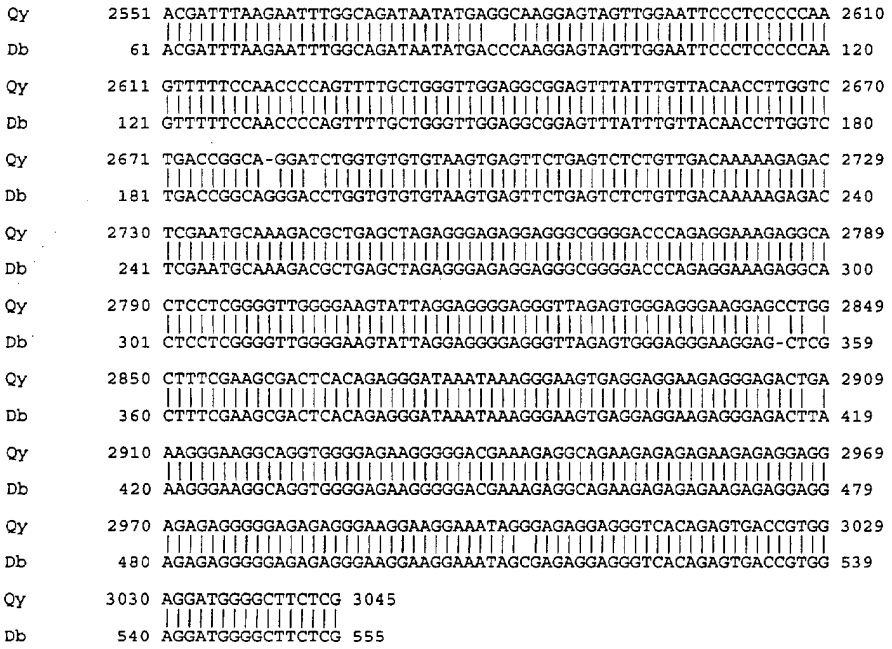
RESULT 5
AAC71827
ID AAC71827 standard; DNA; 839 BP.
XX
AC AAC71827;
XX
DT 09-FEB-2001 (first entry)
XX
DE Single nucleotide polymorphism containing sequence #552.
XX
KW Single nucleotide polymorphism; SNP; human; genetic disease;
KW disease susceptibility; cardiovascular system; endocrine system;
KW neurological system; forensic testing; paternity testing; ds.
XX
OS Homo sapiens.
XX
PN WO200058519-A2.
XX
PD 05-OCT-2000.
XX
PF 30-MAR-2000; 2000WO-US008440.
XX
PR 31-MAR-1999; 99US-0127248P.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PA (AFFY-) AFFYMETRIX INC.
XX
PI Altshuler D, Cargill M, Daley GQ, Ireland JS, Lander ES;
PI Lipshutz RJ, Patil N, Sklar P;
XX
DR WPI; 2000-611722/58.
XX
PT Nucleic acid selected from one of 106 genes comprising single nucleotide
PT polymorphisms, allele-specific oligonucleotides to the genes are useful
PT for phenotypic correlations, forensics, paternity testing, medicine and
PT genetic analysis.
XX
PS Claim 1; Fig 5; 214pp; English.
XX
CC The present invention is concerned with a number of human single
CC nucleotide polymorphisms (SNPs) which the inventors identified in human
CC genes. These SNPs can be used in disease diagnosis and prediction of an
CC individual's susceptibility to disease, in forensic and paternity testing
CC and in genetic mapping. In particular, the SNPs of the invention can be
CC used to diagnose susceptibility to diseases of the cardiovascular,
CC endocrine and neurological systems, such as coronary artery disease,
CC schizophrenia, cancer, autoimmune diseases, Alzheimer's and Parkinson's
CC diseases. Note: The degenerate codon within the sequence represents the
CC position of an SNP, for example the letter S represents a polymorphism
CC where the nucleotide may be C or G
XX
SQ Sequence 839 BP; 212 A; 148 C; 287 G; 191 T; 0 U; 1 Other;

Query Match 24.4%; Score 742.4; DB 3; Length 839;
Best Local Similarity 98.7%; Pred. No. 4.2e-131;
Matches 768; Conservative 1; Mismatches 7; Indels 2; Gaps 2;

```
Qy       2269 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGTTGTTG 2328
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db          1 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAGCGGAGGTACCGTTTTGTTGTTGTTG 60

Qy       2329 TCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGC 2388
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db         61 TCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGYTGGAGTGC 120

Qy       2389 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATC 2448
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATC 180

Qy       2449 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 2508
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 240

Qy       2509 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 2568
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 300

Qy       2569 AGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 2628
             ||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 360

Qy       2629 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGATCTGG 2687
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||||
Db        361 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGACCTGG 420

Qy       2688 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 2747
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 480

Qy       2748 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 2807
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 540

Qy       2808 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGACTCACA 2867
             |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||||||
Db        541 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCGCTTTCGAAGCGACTCACA 599
```

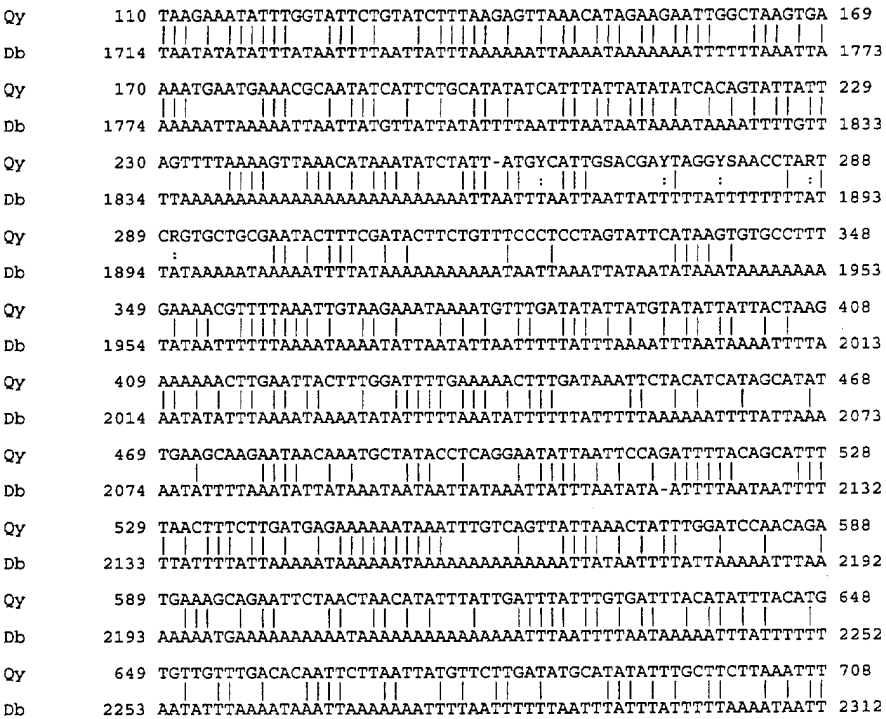
```
Qy       2868 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGGTGGGG 2927
             ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        600 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTAAAGGGAAGGCAGGTGGGG 659

Qy       2928 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 2987
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 719

Qy       2988 AAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 3045
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        720 AAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 777


RESULT 6
AAC71824
ID   AAC71824 standard; DNA; 839 BP.
XX
AC   AAC71824;
XX
DT   09-FEB-2001  (first entry)
XX
DE   Single nucleotide polymorphism containing sequence #551.
XX
KW   Single nucleotide polymorphism; SNP; human; genetic disease;
KW   disease susceptibility; cardiovascular system; endocrine system;
KW   neurological system; forensic testing; paternity testing; ds.
XX
OS   Homo sapiens.
XX
PN   WO200058519-A2.
XX
PD   05-OCT-2000.
XX
PF   30-MAR-2000; 2000WO-US008440.
XX
PR   31-MAR-1999;   99US-0127248P.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (AFFY-) AFFYMETRIX INC.
XX
PI   Altshuler D,  Cargill M,  Daley GQ,  Ireland JS,  Lander ES;
PI   Lipshutz RJ,  Patil N,  Sklar P;
XX
DR   WPI; 2000-611722/58.
XX
PT   Nucleic acid selected from one of 106 genes comprising single nucleotide
PT   polymorphisms, allele-specific oligonucleotides to the genes are useful
PT   for phenotypic correlations, forensics, paternity testing, medicine and
PT   genetic analysis.
XX
PS   Claim 1; Fig 5; 214pp; English.
XX
CC   The present invention is concerned with a number of human single
CC   nucleotide polymorphisms (SNPs) which the inventors identified in human
CC   genes. These SNPs can be used in disease diagnosis and prediction of an
CC   individual's susceptibility to disease, in forensic and paternity testing
CC   and in genetic mapping. In particular, the SNPs of the invention can be
CC   used to diagnose susceptibility to diseases of the cardiovascular,
CC   endocrine and neurological systems, such as coronary artery disease,
CC   schizophrenia, cancer, autoimmune diseases, Alzheimer's and Parkinson's
CC   diseases. Note: The degenerate codon within the sequence represents the
CC   position of an SNP, for example the letter S represents a polymorphism
CC   where the nucleotide may be C or G
XX
SQ   Sequence 839 BP; 212 A; 148 C; 287 G; 191 T; 0 U; 1 Other;

  Query Match           24.4%;  Score 742.4;  DB 3;  Length 839;
  Best Local Similarity  98.7%;  Pred. No. 4.2e-131;
  Matches  768;  Conservative    1;  Mismatches    7;  Indels    2;  Gaps    2;

Qy       2269 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGTTGTTG 2328
             |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          1 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAGCGGAGGTACCGTTTTGTTGTTGTTG 60

Qy       2329 TCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGC 2388
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db         61 TCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTYGCCCAGGCTGGAGTGC 120

Qy       2389 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATC 2448
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATC 180

Qy       2449 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 2508
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 240

Qy       2509 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 2568
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 300

Qy       2569 AGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 2628
             |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||  |
Db        301 AGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 360

Qy       2629 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGATCTGG 2687
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||||
Db        361 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGACCTGG 420

Qy       2688 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 2747
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 480

Qy       2748 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 2807
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 540

Qy       2808 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGACTCACA 2867
             |||||||||||||||||||||||||||||||||||||| || ||||||||||||||||||||
Db        541 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCGCTTTCGAAGCGACTCACA 599

Qy       2868 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGGTGGGG 2927
             ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        600 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTAAAGGGAAGGCAGGTGGGG 659

Qy       2928 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 2987
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 719

Qy       2988 AAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 3045
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        720 AAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 777


RESULT 7
AAC71830
ID   AAC71830 standard; DNA; 839 BP.
XX
AC   AAC71830;
XX
DT   09-FEB-2001  (first entry)
XX
DE   Single nucleotide polymorphism containing sequence #553.
XX
KW   Single nucleotide polymorphism; SNP; human; genetic disease;
KW   disease susceptibility; cardiovascular system; endocrine system;
KW   neurological system; forensic testing; paternity testing; ds.
XX
OS   Homo sapiens.
XX
PN   WO200058519-A2.
XX
PD   05-OCT-2000.
XX
PF   30-MAR-2000; 2000WO-US008440.
XX
```

```
PR   31-MAR-1999;   99US-0127248P.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (AFFY-) AFFYMETRIX INC.
XX
PI   Altshuler D,  Cargill M,  Daley GQ,  Ireland JS,  Lander ES;
PI   Lipshutz RJ,  Patil N,  Sklar P;
XX
DR   WPI; 2000-611722/58.
XX
PT   Nucleic acid selected from one of 106 genes comprising single nucleotide
PT   polymorphisms, allele-specific oligonucleotides to the genes are useful
PT   for phenotypic correlations, forensics, paternity testing, medicine and
PT   genetic analysis.
XX
PS   Claim 1; Fig 5; 214pp; English.
XX
CC   The present invention is concerned with a number of human single
CC   nucleotide polymorphisms (SNPs) which the inventors identified in human
CC   genes. These SNPs can be used in disease diagnosis and prediction of an
CC   individual's susceptibility to disease, in forensic and paternity testing
CC   and in genetic mapping. In particular, the SNPs of the invention can be
CC   used to diagnose susceptibility to diseases of the cardiovascular,
CC   endocrine and neurological systems, such as coronary artery disease,
CC   schizophrenia, cancer, autoimmune diseases, Alzheimer's and Parkinson's
CC   diseases. Note: The degenerate codon within the sequence represents the
CC   position of an SNP, for example the letter S represents a polymorphism
CC   where the nucleotide may be C or G
XX
SQ   Sequence 839 BP; 212 A; 148 C; 287 G; 191 T; 0 U; 1 Other;

  Query Match             24.4%;  Score 742.4;  DB 3;  Length 839;
  Best Local Similarity   98.7%;  Pred. No. 4.2e-131;
  Matches  768;  Conservative    1;  Mismatches    7;  Indels    2;  Gaps    2;

Qy       2269 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGTTGTTG 2328
             |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          1 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAGCGGAGGTACCGTTTTGTTGTTGTTG 60

Qy       2329 TCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGC 2388
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGC 120

Qy       2389 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATC 2448
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        121 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAASGAAGACACACTCGGTCTTCTTCCATC 180

Qy       2449 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 2508
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 240

Qy       2509 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 2568
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 300

Qy       2569 AGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 2628
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 360

Qy       2629 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGATCTGG 2687
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||||
Db        361 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGACCTGG 420

Qy       2688 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 2747
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 480

Qy       2748 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 2807
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 540

Qy       2808 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGACTCACA 2867
             |||||||||||||||||||||||||||||||||| || ||||||||||||||||||||
Db        541 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCGCTTTCGAAGCGACTCACA 599

Qy       2868 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGGTGGGG 2927
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        600 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTAAAGGGAAGGCAGGTGGGG 659

Qy       2928 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 2987
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 719

Qy       2988 AAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 3045
             ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        720 AAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 777


RESULT 8
AAH88477
ID   AAH88477 standard; DNA; 2512 BP.
XX
AC   AAH88477;
XX
DT   26-FEB-2002  (first entry)
XX
DE   CNS disorder-related biallelic marker #11 from 5HT1A gene.
XX
KW   Single nucleotide polymorphism; SNP; biallelic marker; human;
KW   central nervous system disorder; CNS; serotonin receptor; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1501
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200151659-A2.
XX
PD   19-JUL-2001.
XX
PF   11-JAN-2001; 2001WO-IB000116.
XX
PR   13-JAN-2000; 2000US-0175854P.
XX
PA   (GEST ) GENSET.
XX
PI   Chu T,  Blumenfeld M,  Cohen D;
XX
DR   WPI; 2001-483085/52.
XX
PT   Isolated polynucleotides, useful for genotyping nucleic acids for
PT   biallelic markers for the diagnosis of depression, comprises central
PT   nervous system disorder related biallelic marker.
XX
PS   Claim 1; Page 344-345; 519pp; English.
XX
CC   The present invention relates to biallelic markers derived from human
CC   genes involved in central nervous system (CNS) disorders. The present
CC   sequence is one such biallelic marker derived from a human serotonin
CC   receptor gene. This marker has a single nucleotide polymorphism (SNP) and
CC   is useful in determining the genetic predisposition of individuals to CNS
CC   disorders, by identifying the nucleotides at a set of genetic markers in
CC   a biological sample, where the markers comprise at least one CNS disorder
CC   related marker
XX
SQ   Sequence 2512 BP; 537 A; 698 C; 748 G; 522 T; 0 U; 7 Other;

  Query Match             17.2%;  Score 522.4;  DB 4;  Length 2512;
  Best Local Similarity   98.6%;  Pred. No. 2.3e-89;
  Matches  548;  Conservative    0;  Mismatches    6;  Indels    2;  Gaps    2;

Qy       2491 CCCATTCAGGCTCCCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGAC 2550
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCATTCAGGCTCCCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGAC 60
```



RESULT 9
ABZ10246
ID ABZ10246 standard; DNA; 8056 BP.
XX
AC ABZ10246;
XX
DT 16-JAN-2003 (first entry)
XX
DE Haematopoietic cell proliferation disorder related DNA sequence #386.
XX
KW Human; haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200277272-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-EP003401.
XX
PR 26-MAR-2001; 2001US-0278333P.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI Schwope I, Ziebarth H;
XX
DR WPI; 2003-018942/01.
XX
PT Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX
PS Claim 28; SEQ ID NO 386; 117pp; English.
XX
CC The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used: for
CC differentiating between healthy haematopoietic cells and proliferative
CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related DNA
CC sequences. The nucleotide sequences from the present invention can also
CC be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferative disorders. The present method enables a
CC highly specific classification of haematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients
XX
SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;

Query Match 2.9%; Score 89.2; DB 7; Length 8056;
Best Local Similarity 47.0%; Pred. No. 2.9e-07;
Matches 357; Conservative 5; Mismatches 393; Indels 5; Gaps 3;



GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 12:31:35 ; Search time 8774 Seconds
(without alignments)
12211.276 Million cell updates/sec

Title: US-09-430-412A-1
Perfect score: 3045
Sequence: 1 atcatcaataatatccgtta.....gtggaggatggggcttctcg 3045

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US092A_COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US092B_COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US093A_COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US093B_COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US094_COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US095B_COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US095C_COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US095D_COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US096A_COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US096C_COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US096D_COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US096E_COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US097A_COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US098A_COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US098B_COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US098C_COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US098D_COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US099A_COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US099B_COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US099C_COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US099D_COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US099E_COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US099F_COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US100A_COMB.seq:*
44: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
45: /cgn2_6/ptodata/2/pna/US101A_COMB.seq:*
46: /cgn2_6/ptodata/2/pna/US101B_COMB.seq:*
47: /cgn2_6/ptodata/2/pna/US102A_COMB.seq:*
48: /cgn2_6/ptodata/2/pna/US102B_COMB.seq:*
49: /cgn2_6/ptodata/2/pna/US103A_COMB.seq:*
50: /cgn2_6/ptodata/2/pna/US103B_COMB.seq:*
51: /cgn2_6/ptodata/2/pna/US104A_COMB.seq:*
52: /cgn2_6/ptodata/2/pna/US104B_COMB.seq:*
53: /cgn2_6/ptodata/2/pna/US106_COMB.seq:*
54: /cgn2_6/ptodata/2/pna/US107A_COMB.seq:*
55: /cgn2_6/ptodata/2/pna/US107B_COMB.seq:*
56: /cgn2_6/ptodata/2/pna/US6000_COMB.seq:*
57: /cgn2_6/ptodata/2/pna/US6001_COMB.seq:*
58: /cgn2_6/ptodata/2/pna/US6002_COMB.seq:*
59: /cgn2_6/ptodata/2/pna/US6003_COMB.seq:*
60: /cgn2_6/ptodata/2/pna/US6004_COMB.seq:*
61: /cgn2_6/ptodata/2/pna/US6005_COMB.seq:*
62: /cgn2_6/ptodata/2/pna/US6006_COMB.seq:*
63: /cgn2_6/ptodata/2/pna/US6007_COMB.seq:*
64: /cgn2_6/ptodata/2/pna/US6008_COMB.seq:*
65: /cgn2_6/ptodata/2/pna/US6009_COMB.seq:*
66: /cgn2_6/ptodata/2/pna/US6010_COMB.seq:*
67: /cgn2_6/ptodata/2/pna/US6011_COMB.seq:*
68: /cgn2_6/ptodata/2/pna/US6012_COMB.seq:*
69: /cgn2_6/ptodata/2/pna/US6013_COMB.seq:*
70: /cgn2_6/ptodata/2/pna/US6014_COMB.seq:*
71: /cgn2_6/ptodata/2/pna/US6015_COMB.seq:*
72: /cgn2_6/ptodata/2/pna/US6016_COMB.seq:*
73: /cgn2_6/ptodata/2/pna/US6017_COMB.seq:*
74: /cgn2_6/ptodata/2/pna/US6018_COMB.seq:*
75: /cgn2_6/ptodata/2/pna/US6019_COMB.seq:*
76: /cgn2_6/ptodata/2/pna/US6020_COMB.seq:*
77: /cgn2_6/ptodata/2/pna/US6021_COMB.seq:*
78: /cgn2_6/ptodata/2/pna/US6022_COMB.seq:*
79: /cgn2_6/ptodata/2/pna/US6023A_COMB.seq:*
80: /cgn2_6/ptodata/2/pna/US6023B_COMB.seq:*
81: /cgn2_6/ptodata/2/pna/US6024_COMB.seq:*
82: /cgn2_6/ptodata/2/pna/US6025_COMB.seq:*
83: /cgn2_6/ptodata/2/pna/US6026_COMB.seq:*
84: /cgn2_6/ptodata/2/pna/US6027_COMB.seq:*
85: /cgn2_6/ptodata/2/pna/US6028_COMB.seq:*
86: /cgn2_6/ptodata/2/pna/US6029_COMB.seq:*
87: /cgn2_6/ptodata/2/pna/US6030_COMB.seq:*
88: /cgn2_6/ptodata/2/pna/US6031_COMB.seq:*
89: /cgn2_6/ptodata/2/pna/US6032_COMB.seq:*
90: /cgn2_6/ptodata/2/pna/US6033_COMB.seq:*
91: /cgn2_6/ptodata/2/pna/US6034_COMB.seq:*
92: /cgn2_6/ptodata/2/pna/US6035_COMB.seq:*
93: /cgn2_6/ptodata/2/pna/US6036_COMB.seq:*
94: /cgn2_6/ptodata/2/pna/US6037_COMB.seq:*
95: /cgn2_6/ptodata/2/pna/US6038_COMB.seq:*
96: /cgn2_6/ptodata/2/pna/US6039_COMB.seq:*
97: /cgn2_6/ptodata/2/pna/US6040_COMB.seq:*
98: /cgn2_6/ptodata/2/pna/US6041_COMB.seq:*
99: /cgn2_6/ptodata/2/pna/US6042_COMB.seq:*
100: /cgn2_6/ptodata/2/pna/US6043_COMB.seq:*
101: /cgn2_6/ptodata/2/pna/US6044_COMB.seq:*
102: /cgn2_6/ptodata/2/pna/US6045_COMB.seq:*
103: /cgn2_6/ptodata/2/pna/US6046_COMB.seq:*
104: /cgn2_6/ptodata/2/pna/US6047_COMB.seq:*
105: /cgn2_6/ptodata/2/pna/US6048_COMB.seq:*
106: /cgn2_6/ptodata/2/pna/US6049_COMB.seq:*
107: /cgn2_6/ptodata/2/pna/US6050_COMB.seq:*
108: /cgn2_6/ptodata/2/pna/US6051_COMB.seq:*
109: /cgn2_6/ptodata/2/pna/US6052_COMB.seq:*
110: /cgn2_6/ptodata/2/pna/US6053_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3042.2	99.9	3045	20	US-09-430-412A-1	Sequence 1, Appli
2	2731	89.7	13269	103	US-60-466-412-83506	Sequence 83506, A
3	2731	89.7	14011	107	US-60-500-315-11863	Sequence 11863, A
4	2730.6	89.7	13482	107	US-60-500-315-11364	Sequence 11364, A
5	1381.4	45.4	2112	76	US-60-205-423-94	Sequence 94, Appl
6	747.8	24.6	1204	1	PCT-US00-40519-1	Sequence 1, Appli
7	747.8	24.6	1204	43	US-10-049-407-1	Sequence 1, Appli
8	747.8	24.6	2722	1	PCT-US00-40519-3	Sequence 3, Appli
9	747.8	24.6	2722	43	US-10-049-407-3	Sequence 3, Appli
10	742.4	24.4	839	23	US-09-541-946-2477	Sequence 2477, Ap
11	742.4	24.4	839	23	US-09-541-946-2478	Sequence 2478, Ap
12	742.4	24.4	839	23	US-09-541-946-2479	Sequence 2479, Ap
c 13	428.2	14.1	559	26	US-09-634-306B-196186	Sequence 196186,
c 14	428.2	14.1	559	26	US-09-634-306B-196187	Sequence 196187,
c 15	428.2	14.1	559	43	US-10-027-632-196186	Sequence 196186,
c 16	428.2	14.1	559	43	US-10-027-632-196187	Sequence 196187,
17	361.2	11.9	378	21	US-09-528-409-45109	Sequence 45109, A
18	361.2	11.9	378	38	US-09-933-524-45109	Sequence 45109, A
19	361.2	11.9	378	38	US-09-933-524A-45109	Sequence 45109, A
20	320.2	10.5	452	21	US-09-528-409-48578	Sequence 48578, A
21	320.2	10.5	452	38	US-09-933-524-48578	Sequence 48578, A
22	320.2	10.5	452	38	US-09-933-524A-48578	Sequence 48578, A
c 23	194.8	6.4	241	23	US-09-540-229-134492	Sequence 134492,
24	174.6	5.7	201	107	US-60-500-315-21639	Sequence 21639, A
c 25	125.6	4.1	7218	8	US-08-466-194-14	Sequence 14, Appl
c 26	95	3.1	1275	48	US-10-266-090-36380	Sequence 36380, A
27	89.2	2.9	8056	52	US-10-473-126-386	Sequence 386, App
28	88.4	2.9	1111	76	US-60-207-458-37399	Sequence 37399, A
29	86.8	2.9	8056	52	US-10-473-126-240	Sequence 240, App
c 30	86.4	2.8	970	48	US-10-266-090-30537	Sequence 30537, A
31	85.8	2.8	192014	81	US-60-248-823-33	Sequence 33, Appl
32	85.6	2.8	144759	78	US-60-226-176-511	Sequence 511, App
33	85.6	2.8	144759	79	US-60-233-468-511	Sequence 511, App
34	85.6	2.8	144759	88	US-60-313-371-511	Sequence 511, App
c 35	85.4	2.8	1326	28	US-09-663-779-8205	Sequence 8205, Ap
c 36	85	2.8	618	20	US-09-404-520-13571	Sequence 13571, A
c 37	83.6	2.7	824	76	US-60-207-458-18907	Sequence 18907, A
c 38	83.6	2.7	2453797	39	US-09-947-911-331	Sequence 331, App
39	83.4	2.7	1040	48	US-10-266-090-12299	Sequence 12299, A
40	83.2	2.7	11745	47	US-10-240-453-206	Sequence 206, App
c 41	83	2.7	742	28	US-09-663-779-3024	Sequence 3024, Ap
c 42	83	2.7	1140	48	US-10-266-090-19455	Sequence 19455, A
c 43	83	2.7	187169	39	US-09-949-016-12776	Sequence 12776, A
c 44	83	2.7	191569	39	US-09-949-016-15940	Sequence 15940, A
c 45	82.8	2.7	1107	48	US-10-266-090-23011	Sequence 23011, A

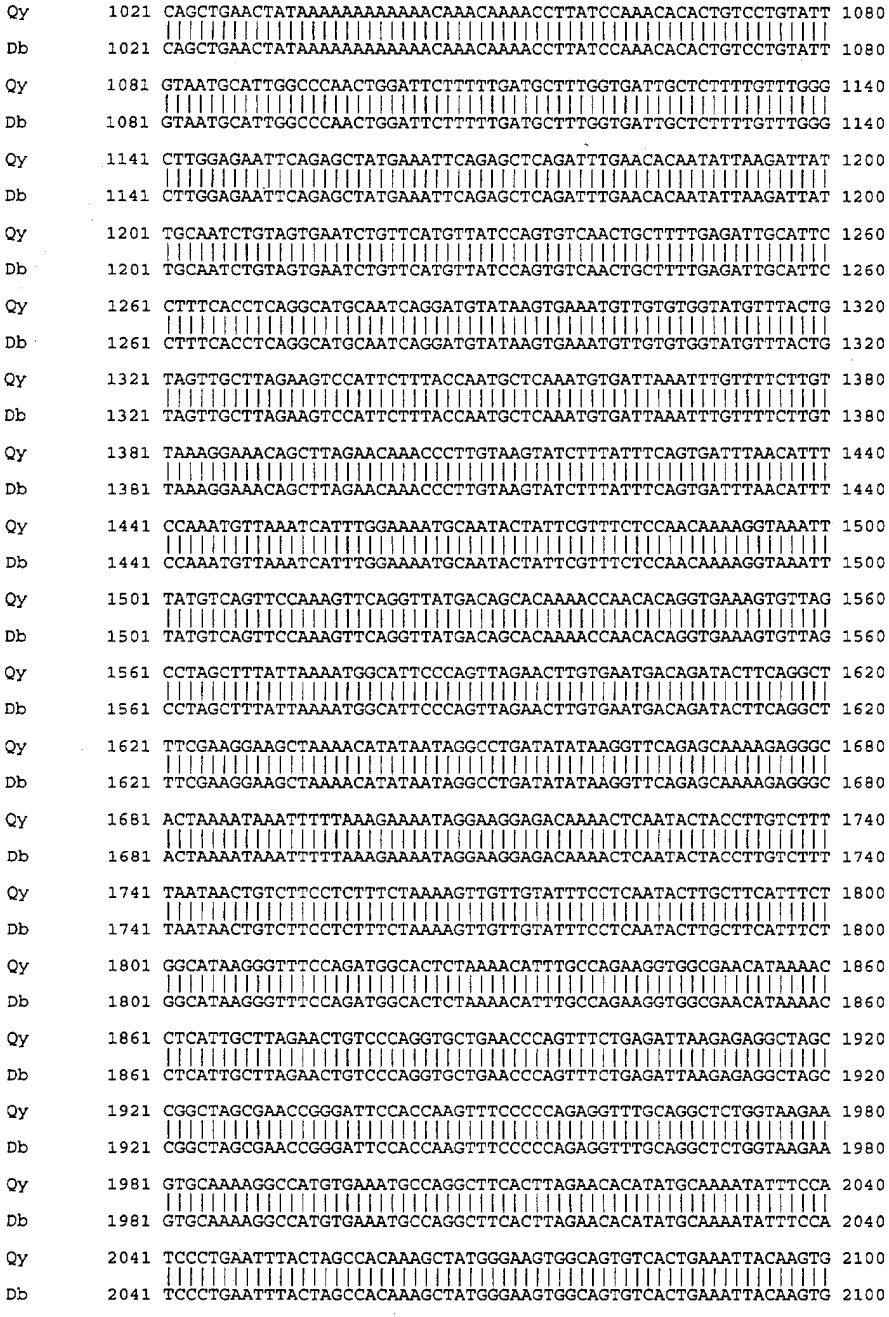
ALIGNMENTS

```
RESULT 1
US-09-430-412A-1
; Sequence 1, Application US/09430412A
; GENERAL INFORMATION:
;  APPLICANT: Albert Paul; and Lemonde, Sylvie
;  TITLE OF INVENTION: Mutations of the 5' region of the human 5-HT1A gene,
;  TITLE OF INVENTION:  associated proteins of the 5' region and a diagnostic
;  TITLE OF INVENTION:  test for major depression and related mental illnesses
;  FILE REFERENCE: 881014US1
;  CURRENT APPLICATION NUMBER: US/09/430,412A
;  CURRENT FILING DATE:  1999-10-29
;  PRIOR APPLICATION NUMBER: 60/106,375
;  PRIOR FILING DATE: 1998-10-30
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 3045
;   TYPE: DNA
;   ORGANISM: human
```

US-09-430-412A-1

Query Match 99.9%; Score 3042.2; DB 20; Length 3045;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3045; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATCATCAATAATATCCGTTATAAAGCTTGCTTTTCTTTAGGTTAACTTTAGAGGCCTTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCATCAATAATATCCGTTATAAAGCTTGCTTTTCTTTAGGTTAACTTTAGAGGCCTTGA 60

Qy         61 AGAATAAGAGCTCATCTCTTTACAGGAGCTTTGGTTTGCAGCATTTACTTAAGAAATATT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAATAAGAGCTCATCTCTTTACAGGAGCTTTGGTTTGCAGCATTTACTTAAGAAATATT 120

Qy        121 TGGTATTCTGTATCTTTAAGAGTTAAACATAGAAGAATTGGCTAAGTGAAAATGAATGAA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGTATTCTGTATCTTTAAGAGTTAAACATAGAAGAATTGGCTAAGTGAAAATGAATGAA 180

Qy        181 ACGCAATATCATTCTGCATATATCATTTATTATATATCACAGTATTATTAGTTTTAAAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACGCAATATCATTCTGCATATATCATTTATTATATATCACAGTATTATTAGTTTTAAAAG 240

Qy        241 TTAAACATAAATATCTATTATGYCATTGSACGAYTAGGYSAACCTARTCRGTGCTGCGAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTAAACATAAATATCTATTATGYCATTGSACGAYTAGGYSAACCTARTCRGTGCTGCGAA 300

Qy        301 TACTTTCGATACTTCTGTTTCCCTCCTAGTATTCATAAGTGTGCCTTTGAAAACGTTTTA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TACTTTCGATACTTCTGTTTCCCTCCTAGTATTCATAAGTGTGCCTTTGAAAACGTTTTA 360

Qy        361 AATTGTAAGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACTTGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AATTGTAAGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACTTGAA 420

Qy        421 TTACTTTGGATTTTGAAAAACTTTGATAAATTCTACATCATAGCATATTGAAGCAAGAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTACTTTGGATTTTGAAAAACTTTGATAAATTCTACATCATAGCATATTGAAGCAAGAAT 480

Qy        481 AACAAATGCTATACCTCAGGAATATTAATTCCAGATTTTACAGCATTTTAACTTTCTTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACAAATGCTATACCTCAGGAATATTAATTCCAGATTTTACAGCATTTTAACTTTCTTGA 540

Qy        541 TGAGAAAAAATAAATTTGTCAGTTATTAAACTATTTGGATCCAACAGATGAAAGCAGAAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGAGAAAAAATAAATTTGTCAGTTATTAAACTATTTGGATCCAACAGATGAAAGCAGAAT 600

Qy        601 TCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTTGACA 660

Qy        661 CAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTCCTTT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTCCTTT 720

Qy        721 ATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTCAGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTCAGCC 780

Qy        781 TTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTTAACTGTAAATAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTTAACTGTAAATAT 840

Qy        841 AGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTTAAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTTAAAA 900

Qy        901 GACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTGCATGCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTGCATGCC 960

Qy        961 ATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTGAATGATTGAGTGGGACTGTGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTGAATGATTGAGTGGGACTGTGC 1020
```



RESULT 2
US-60-466-412-83506
; Sequence 83506, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF

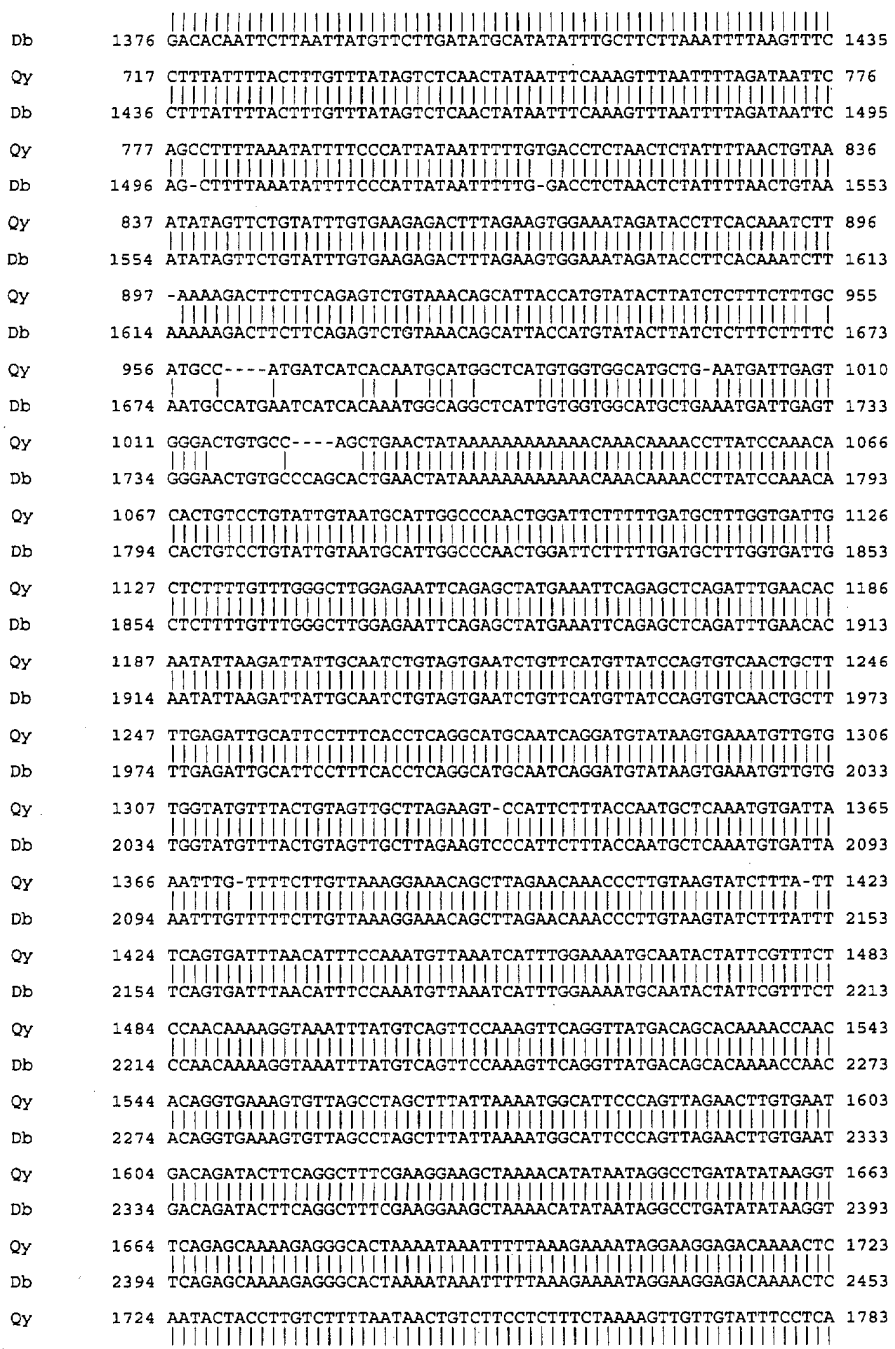
```
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 83506
;  LENGTH: 13269
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-60-466-412-83506

 Query Match            89.7%;  Score 2731;  DB 103;  Length 13269;
 Best Local Similarity  97.4%;  Pred. No. 0;
 Matches 2984;  Conservative    5;  Mismatches   45;  Indels   29;  Gaps   20;

Qy          1 ATCATCAATAATATCCGTTATAAAGCTTGCTTTTCTTTAGGTTAACTTTAGAGGCCTTGA 60
Db       2554 ATCATCAATAATATCCGTTATAAAGCTTGCTTTTCTTTAGGTTAACTTTAGAGGCCTTGA 2613

Qy         61 AGAATAAGAGCTCATCTCTTTA--CAGGAGCTTTGGTTTGCAGCATTTACTTAAGAAATA 118
Db       2614 AGAATAAGAGCTCATCTCTTTATGCAGGAGCTTTGGTTTGCAGCATTTACTTAAGAAATA 2673

Qy        119 TTTGGTATTCTGTATCTTTAAGAGTTAAACATAGAAGAATTGGCTAAGTGAAAATGAATG 178
Db       2674 TTTGGTATTCTGTATCTTTAAGAGTTAAACATAGAAGAATTGGCTAAGTGAAAATGAA-G 2732

Qy        179 AAACGCAATATCATTCTGCATATATCATTTATTATATATCACAGTATTATTAGTTTT-AA 237
Db       2733 AAACGCAATATCATTCTGCATATATCATTTATTATATATCACAGTATTATTAGTTTTAAA 2792

Qy        238 AAGTTAAACATAAATATCTATTATGYCATTGSACGAYTAGGYSAACCTARTCRGTGCTGC 297
Db       2793 AAGTTAAACATAAATATCTATTA-GTCACTTGCAGACTATGGAACCATAGTCAGTGCTGC 2851

Qy        298 GAATACTTTCGATACTTCTGTTTCCCTCCTAGTATTCATAAGTGTGCCTTTGAAAACGTT 357
Db       2852 ATA---CTTTCATACTTCTGTTTCCCTCCTAGTATTCATAAGTGTGCCTTTAAAAAC-TT 2907

Qy        358 TTAAATTGTAAGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACTT 417
Db       2908 TTAAATTGTAAGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACTT 2967

Qy        418 GAATTACTTTGGATTTTGAAAAACTTTGATAAATTCTACATCATAGCATATTGAAGCAAG 477
Db       2968 GAATTACTTTGGATTTTTAAAAACTTTGATAAATTCTACATCATAGCATATTGAAGC-AG 3026

Qy        478 AATAACAAATGCTATACCTCAGGAATATTAATTCCAGATTTTACAGCATTTTAACTTTCT 537
Db       3027 AATAACAAATGCTATACCTCAGGAATATTAATTCCAGATTTTACAGCATTTTAACTTTCT 3086

Qy        538 TGATGAGAAAAAATAAATTTGTCAGTTATTAAACTATTT-GGATCCAACAGATGAAAGCA 596
Db       3087 TGATGAGAAAAAATAAATTTGTCAGTTATTAAACTATTTGGGATCCAACAGATGAAAGCA 3146

Qy        597 GAATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTT 656
Db       3147 GAATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTT 3206

Qy        657 GACACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTC 716
Db       3207 GACACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTC 3266

Qy        717 CTTTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTC 776
Db       3267 CTTTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTC 3326

Qy        777 AGCCTTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTTAACTGTAA 836
Db       3327 AG-CTTTTAAATATTTTCCCATTATAATTTTTG-GACCTCTAACTCTATTTTAACTGTAA 3384

Qy        837 ATATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTT 896
Db       3385 ATATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTT 3444

Qy        897 -AAAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTGC 955
Db       3445 AAAAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTTC 3504

Qy        956 ATGCC----ATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTG-AATGATTGAGT 1010
Db       3505 AATGCCATGAATCATCACAAATGGCAGGCTCATTGTGGTGGCATGCTGAAATGATTGAGT 3564

Qy       1011 GGGACTGTGCC----AGCTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACA 1066
Db       3565 GGGAACTGTGCCCAGCACTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACA 3624

Qy       1067 CACTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTG 1126
Db       3625 CACTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTG 3684

Qy       1127 CTCTTTTGTTTGGGCTTGGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACAC 1186
Db       3685 CTCTTTTGTTTGGGCTTGGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACAC 3744

Qy       1187 AATATTAAGATTATTGCAATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTT 1246
Db       3745 AATATTAAGATTATTGCAATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTT 3804

Qy       1247 TTGAGATTGCATTCCTTTCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTGTG 1306
Db       3805 TTGAGATTGCATTCCTTTCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTGTG 3864

Qy       1307 TGGTATGTTTACTGTAGTTGCTTAGAAGT-CCATTCTTTACCAATGCTCAAATGTGATTA 1365
Db       3865 TGGTATGTTTACTGTAGTTGCTTAGAAGTCCCATTCTTTACCAATGCTCAAATGTGATTA 3924

Qy       1366 AATTTG-TTTTCTTGTTAAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTTA-TT 1423
Db       3925 AATTTGTTTTTCTTGTTAAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTTATTT 3984

Qy       1424 TCAGTGATTTAACATTTCCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTTTCT 1483
Db       3985 TCAGTGATTTAACATTTCCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTTTCT 4044

Qy       1484 CCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACCAAC 1543
Db       4045 CCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACCAAC 4104

Qy       1544 ACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTGAAT 1603
Db       4105 ACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTGAAT 4164

Qy       1604 GACAGATACTTCAGGCTTTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAAGGT 1663
Db       4165 GACAGATACTTCAGGCTTTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAAGGT 4224

Qy       1664 TCAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAACTC 1723
Db       4225 TCAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAACTC 4284

Qy       1724 AATACTACCTTGTCTTTTAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCA 1783
Db       4285 AATACTACCTTGTCTTTTAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCA 4344

Qy       1784 ATACTTGCTTCATTTCTGGCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAA 1843
Db       4345 ATACTTGCTTCATTTCT-GCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAA 4403

Qy       1844 GGTGGCGAACATAAAACCTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGA 1903
Db       4404 GGTGGCG-ACATAAAACCTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGA 4462

Qy       1904 GATTAAGAGAGGCTAGCCGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTT 1963
Db       4463 GATTAAGAGAGGCTAGCCGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTT 4522
```

```
Qy       1964 GCAGGCTCTGGTAAGAAGTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACA 2023
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4523 GCAGGCTCTGGTAAGAAGTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACA 4582

Qy       2024 TATGCAAAATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGT 2083
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4583 TATGCAAAATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGT 4642

Qy       2084 CACTGAAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCA 2143
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4643 CACTGAAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCA 4702

Qy       2144 CACTGAGTTTTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACAC 2203
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4703 CACTGAGTTTTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACAC 4762

Qy       2204 GTCCTTTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAAGTCACAGGCAAT 2263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4763 GTCCTTTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAGTCACAGGCAAT 4822

Qy       2264 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 2323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4823 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 4882

Qy       2324 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 2383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4883 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 4942

Qy       2384 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 2443
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       4943 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAAGGAAGACACACTCGGTCTTCTT 5002

Qy       2444 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 2503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5003 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 5062

Qy       2504 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 2563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5063 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 5122

Qy       2564 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 2623
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5123 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 5182

Qy       2624 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGA 2682
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       5183 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGA 5242

Qy       2683 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5243 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 5302

Qy       2743 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 2802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5303 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 5362

Qy       2803 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5363 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 5422

Qy       2863 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 2922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5423 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 5482

Qy       2923 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 2982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5483 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 5542

Qy       2983 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 3042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5543 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 5602

Qy       3043 TCG 3045
              |||
Db       5603 TCG 5605
```

RESULT 3
US-60-500-315-11863
; Sequence 11863, Application US/60500315
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN G-PROTEIN COUPLED RECEPTOR PROTEINS, METHODS
; TITLE OF INVENTION: OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001484
; CURRENT APPLICATION NUMBER: US/60/500,315
; CURRENT FILING DATE: 2003-09-05
; NUMBER OF SEQ ID NOS: 69978
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11863
; LENGTH: 14011
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-500-315-11863

Query Match 89.7%; Score 2731; DB 107; Length 14011;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 2984; Conservative 5; Mismatches 45; Indels 29; Gaps 20;

```
Qy          1 ATCATCAATAATATCCGTTATAAAGCTTGCTTTTCTTTAGGTTAACTTTAGAGGCCTTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 ATCATCAATAATATCCGTTATAAAGCTTGCTTTTCTTTAGGTTAACTTTAGAGGCCTTGA 782

Qy         61 AGAATAAGAGCTCATCTCTTTA--CAGGAGCTTTGGTTTGCAGCATTTACTTAAGAAATA 118
              |||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        783 AGAATAAGAGCTCATCTCTTTATGCAGGAGCTTTGGTTTGCAGCATTTACTTAAGAAATA 842

Qy        119 TTTGGTATTCTGTATCTTTAAGAGTTAAACATAGAAGAATTGGCTAAGTGAAAATGAATG 178
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        843 TTTGGTATTCTGTATCTTTAAGAGTTAAACATAGAAGAATTGGCTAAGTGAAAATGAA-G 901

Qy        179 AAACGCAATATCATTCTGCATATATCATTTATTATATATCACAGTATTATTAGTTTT-AA 237
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        902 AAACGCAATATCATTCTGCATATATCATTTATTATATATCACAGTATTATTAGTTTTAAA 961

Qy        238 AAGTTAAACATAAATATCTATTATGYCATTGSACGAYTAGGYSAACCTARTCRGTGCTGC 297
              |||||||||||||||||||||| |:|| | :  ||:|| |  | | ||:||:|||||||
Db        962 AAGTTAAACATAAATATCTATTA-GTCACTTGCAGACTATGGAACCATAGTCAGTGCTGC 1020

Qy        298 GAATACTTTCGATACTTCTGTTTCCCTCCTAGTATTCATAAGTGTGCCTTTGAAAACGTT 357
              |    ||  |||||||||||||||||||||||||||||||||||||||||| |||| ||
Db       1021 ATA---CTTTCATACTTCTGTTTCCCTCCTAGTATTCATAAGTGTGCCTTTAAAAAC-TT 1076

Qy        358 TTAAATTGTAAGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACTT 417
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 TTAAATTGTAAGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACTT 1136

Qy        418 GAATTACTTTGGATTTTGAAAAACTTTGATAAATTCTACATCATAGCATATTGAAGCAAG 477
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||| ||
Db       1137 GAATTACTTTGGATTTTTAAAAACTTTGATAAATTCTACATCATAGCATATTGAAGC-AG 1195

Qy        478 AATAACAAATGCTATACCTCAGGAATATTAATTCCAGATTTTACAGCATTTTAACTTTCT 537
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1196 AATAACAAATGCTATACCTCAGGAATATTAATTCCAGATTTTACAGCATTTTAACTTTCT 1255

Qy        538 TGATGAGAAAAAAATAAATTTGTCAGTTATTAAACTATTT-GGATCCAACAGATGAAAGCA 596
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       1256 TGATGAGAAAAAATAAATTTGTCAGTTATTAAACTATTTGGGATCCAACAGATGAAAGCA 1315

Qy        597 GAATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTT 656
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1316 GAATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTT 1375

Qy        657 GACACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTC 716
```


Db 1376 GACACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTC 1435
Qy 717 CTTTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTC 776
Db 1436 CTTTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTC 1495
Qy 777 AGCCTTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTTAACTGTAA 836
Db 1496 AG-CTTTTAAATATTTTCCCATTATAATTTTTG-GACCTCTAACTCTATTTTAACTGTAA 1553
Qy 837 ATATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTT 896
Db 1554 ATATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTT 1613
Qy 897 -AAAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTGC 955
Db 1614 AAAAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTTC 1673
Qy 956 ATGCC----ATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTG-AATGATTGAGT 1010
Db 1674 AATGCCATGAATCATCACAAATGGCAGGCTCATTGTGGTGGCATGCTGAAATGATTGAGT 1733
Qy 1011 GGGACTGTGCC----AGCTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACA 1066
Db 1734 GGGAACTGTGCCCAGCACTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACA 1793
Qy 1067 CACTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTG 1126
Db 1794 CACTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTG 1853
Qy 1127 CTCTTTTGTTTGGGCTTGGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACAC 1186
Db 1854 CTCTTTTGTTTGGGCTTGGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACAC 1913
Qy 1187 AATATTAAGATTATTGCAATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTT 1246
Db 1914 AATATTAAGATTATTGCAATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTT 1973
Qy 1247 TTGAGATTGCATTCCTTTCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTGTG 1306
Db 1974 TTGAGATTGCATTCCTTTCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTGTG 2033
Qy 1307 TGGTATGTTTACTGTAGTTGCTTAGAAGT-CCATTCTTTACCAATGCTCAAATGTGATTA 1365
Db 2034 TGGTATGTTTACTGTAGTTGCTTAGAAGTCCCATTCTTTACCAATGCTCAAATGTGATTA 2093
Qy 1366 AATTTG-TTTTCTTGTTAAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTTA-TT 1423
Db 2094 AATTTGTTTTTCTTGTTAAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTTATTT 2153
Qy 1424 TCAGTGATTTAACATTTCCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTTTCT 1483
Db 2154 TCAGTGATTTAACATTTCCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTTTCT 2213
Qy 1484 CCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACCAAC 1543
Db 2214 CCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACCAAC 2273
Qy 1544 ACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTGAAT 1603
Db 2274 ACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTGAAT 2333
Qy 1604 GACAGATACTTCAGGCTTTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAAGGT 1663
Db 2334 GACAGATACTTCAGGCTTTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAAGGT 2393
Qy 1664 TCAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAACTC 1723
Db 2394 TCAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAACTC 2453
Qy 1724 AATACTACCTTGTCTTTTAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCA 1783
Db 2454 AATACTACCTTGTCTTTTAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCA 2513
Qy 1784 ATACTTGCTTCATTTCTGGCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAA 1843
Db 2514 ATACTTGCTTCATTTCT-GCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAA 2572
Qy 1844 GGTGGCGAACATAAAACCTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGA 1903
Db 2573 GGTGGCG-ACATAAAACCTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGA 2631
Qy 1904 GATTAAGAGAGGCTAGCCGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTT 1963
Db 2632 GATTAAGAGAGGCTAGCCGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTT 2691
Qy 1964 GCAGGCTCTGGTAAGAAGTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACA 2023
Db 2692 GCAGGCTCTGGTAAGAAGTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACA 2751
Qy 2024 TATGCAAAATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGT 2083
Db 2752 TATGCAAAATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGT 2811
Qy 2084 CACTGAAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCA 2143
Db 2812 CACTGAAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCA 2871
Qy 2144 CACTGAGTTTTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACAC 2203
Db 2872 CACTGAGTTTTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACAC 2931
Qy 2204 GTCCTTTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAGTCACAGGCAAT 2263
Db 2932 GTCCTTTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAGTCACAGGCAAT 2991
Qy 2264 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 2323
Db 2992 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 3051
Qy 2324 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 2383
Db 3052 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 3111
Qy 2384 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 2443
Db 3112 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAAGGAAGACACACTCGGTCTTCTT 3171
Qy 2444 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 2503
Db 3172 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 3231
Qy 2504 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 2563
Db 3232 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 3291
Qy 2564 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 2623
Db 3292 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 3351
Qy 2624 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGA 2682
Db 3352 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGA 3411
Qy 2683 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 2742
Db 3412 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 3471
Qy 2743 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 2802
Db 3472 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 3531
Qy 2803 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 2862
Db 3532 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 3591

```
Qy       2863 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 2922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3592 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 3651

Qy       2923 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 2982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3652 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 3711

Qy       2983 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 3042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3712 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 3771

Qy       3043 TCG 3045
              |||
Db       3772 TCG 3774

RESULT 4
US-60-500-315-11364
; Sequence 11364, Application US/60500315
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
;  TITLE OF INVENTION:  ENCODING HUMAN G-PROTEIN COUPLED RECEPTOR PROTEINS, METHODS
;  TITLE OF INVENTION:  OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001484
;  CURRENT APPLICATION NUMBER: US/60/500,315
;  CURRENT FILING DATE:  2003-09-05
;  NUMBER OF SEQ ID NOS: 69978
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11364
;   LENGTH: 13482
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-500-315-11364

  Query Match             89.7%;  Score 2730.6;  DB 107;  Length 13482;
  Best Local Similarity   97.4%;  Pred. No. 0;
  Matches 2983;  Conservative    6;  Mismatches   45;  Indels   29;  Gaps   20;

Qy          1 ATCATCAATAATATCCGTTATAAAGCTTGCTTTTCTTTAGGTTAACTTTAGAGGCCTTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2767 ATCATCAATAATATCCGTTATAAAGCTTGCTTTTCTTTAGGTTAACTTTAGAGGCCTTGA 2826

Qy         61 AGAATAAGAGCTCATCTCTTTA--CAGGAGCTTTGGTTTGCAGCATTTACTTAAGAAATA 118
              ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db       2827 AGAATAAGAGCTCATCTCTTTATGCAGGAGCTTTGGTTTGCAGCATTTACTTAAGAAATA 2886

Qy        119 TTTGGTATTCTGTATCTTTAAGAGTTAAACATAGAAGAATTGGCTAAGTGAAAATGAATG 178
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       2887 TTTGGTATTCTGTATCTTTAAGAGTTAAACATAGAAGAATTGGCTAAGTGAAAATGAA-G 2945

Qy        179 AAACGCAATATCATTCTGCATATATCATTTATTATATATCACAGTATTATTAGTTTT-AA 237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       2946 AAACGCAATATCATTCTGCATATATCATTTATTATATATCACAGTATTATTAGTTTTAAA 3005

Qy        238 AAGTTAAACATAAATATCTATTATGYCATTGSACGAYTAGGYSAACCTARTCRGTGCTGC 297
              ||||||||||||||||||||||| |:|| | :  ||:|| |  | | ||:||:|||||||
Db       3006 AAGTTAAACATAAATATCTATTA-GTCACTTGCAGACTATGGAACCATAGTCAGTGCTGC 3064

Qy        298 GAATACTTTCGATACTTCTGTTTCCCTCCTAGTATTCATAAGTGTGCCTTTGAAAACGTT 357
                |    ||  |||||||||||||||||||||||||||||||||||||||| ||||| ||
Db       3065 ATA---CTTTCATACTTCTGTTTCCCTCCTAGTATTCATAAGTGTGCCTTTAAAAAC-TT 3120

Qy        358 TTAAATTGTAAGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACTT 417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TTAAATTGTAAGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACTT 3180

Qy        418 GAATTACTTTGGATTTTGAAAAACTTTGATAAATTCTACATCATAGCATATTGAAGCAAG 477
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||| ||
Db       3181 GAATTACTTTGGATTTTTAAAAACTTTGATAAATTCTACATCATAGCATATTGAAGC-AG 3239

Qy        478 AATAACAAATGCTATACCTCAGGAATATTAATTCCAGATTTTACAGCATTTTAACTTTCT 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3240 AATAACAAATGCTATACCTCAGGAATATTAATTCCAGATTTTACAGCATTTTAACTTTCT 3299

Qy        538 TGATGAGAAAAAATAAATTTGTCAGTTATTAAACTATTT-GGATCCAACAGATGAAAGCA 596
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       3300 TGATGAGAAAAAATAAATTTGTCAGTTATTAAACTATTTGGGATCCAACAGATGAAAGCA 3359

Qy        597 GAATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTT 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3360 GAATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTT 3419

Qy        657 GACACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTC 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3420 GACACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTC 3479

Qy        717 CTTTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTC 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3480 CTTTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTC 3539

Qy        777 AGCCTTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTTAACTGTAA 836
              || |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       3540 AG-CTTTTAAATATTTTCCCATTATAATTTTTG-GACCTCTAACTCTATTTTAACTGTAA 3597

Qy        837 ATATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTT 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3598 ATATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTT 3657

Qy        897 -AAAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTGC 955
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       3658 AAAAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTTC 3717

Qy        956 ATGCC----ATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTG-AATGATTGAGT 1010
              |   |    |       || |  ||| |     |||||||||||||| |||||||||||
Db       3718 AATGCCATGAATCATCACAAATGGCAGGCTCATTGTGGTGGCATGCTGAAATGATTGAGT 3777

Qy       1011 GGACTGTGCC----AGCTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACA 1066
              ||||     |      ||||||||||||||||||||||||||||||||||||||||||||
Db       3778 GGGAACTGTGCCCAGCACTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACA 3837

Qy       1067 CACTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTG 1126
              |||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       3838 CACTGTCCTGTATTGTAATGCATTGGCCCAASTGGATTCTTTTTGATGCTTTGGTGATTG 3897

Qy       1127 CTCTTTTGTTTGGGCTTGGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACAC 1186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3898 CTCTTTTGTTTGGGCTTGGAGAATTCAGAGCTATGAAATTCAGAGCTCAGATTTGAACAC 3957

Qy       1187 AATATTAAGATTATTGCAATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTT 1246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3958 AATATTAAGATTATTGCAATCTGTAGTGAATCTGTTCATGTTATCCAGTGTCAACTGCTT 4017

Qy       1247 TTGAGATTGCATTCCTTTCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTGTG 1306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4018 TTGAGATTGCATTCCTTTCACCTCAGGCATGCAATCAGGATGTATAAGTGAAATGTTGTG 4077

Qy       1307 TGGTATGTTTACTGTAGTTGCTTAGAAGT-CCATTCTTTACCAATGCTCAAATGTGATTA 1365
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       4078 TGGTATGTTTACTGTAGTTGCTTAGAAGTCCCATTCTTTACCAATGCTCAAATGTGATTA 4137

Qy       1366 AATTTG-TTTTCTTGTTAAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTTA-TT 1423
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       4138 AATTTGTTTTTCTTGTTAAAGGAAACAGCTTAGAACAAACCCTTGTAAGTATCTTTATTT 4197

Qy       1424 TCAGTGATTTAACATTTCCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTTTCT 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4198 TCAGTGATTTAACATTTCCAAATGTTAAATCATTTGGAAAATGCAATACTATTCGTTTCT 4257

Qy       1484 CCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACCAAC 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4258 CCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTCAGGTTATGACAGCACAAAACCAAC 4317

Qy       1544 ACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTGAAT 1603
```

```
Db       4318 ACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGGCATTCCCAGTTAGAACTTGTGAAT 4377

Qy       1604 GACAGATACTTCAGGCTTTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAAGGT 1663

Db       4378 GACAGATACTTCAGGCTTTCGAAGGAAGCTAAAACATATAATAGGCCTGATATATAAGGT 4437

Qy       1664 TCAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAACTC 1723

Db       4438 TCAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATAGGAAGGAGACAAAACTC 4497

Qy       1724 AATACTACCTTGTCTTTTAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCA 1783

Db       4498 AATACTACCTTGTCTTTTAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCA 4557

Qy       1784 ATACTTGCTTCATTTCTGGCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAA 1843

Db       4558 ATACTTGCTTCATTTCT-GCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAA 4616

Qy       1844 GGTGGCGAACATAAAACCTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGA 1903

Db       4617 GGTGGCG-ACATAAAACCTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGA 4675

Qy       1904 GATTAAGAGAGGCTAGCCGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTT 1963

Db       4676 GATTAAGAGAGGCTAGCCGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTT 4735

Qy       1964 GCAGGCTCTGGTAAGAAGTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACA 2023

Db       4736 GCAGGCTCTGGTAAGAAGTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACA 4795

Qy       2024 TATGCAAAATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGT 2083

Db       4796 TATGCAAAATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGT 4855

Qy       2084 CACTGAAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCA 2143

Db       4856 CACTGAAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCA 4915

Qy       2144 CACTGAGTTTTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACAC 2203

Db       4916 CACTGAGTTTTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACAC 4975

Qy       2204 GTCCTTTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAGTCACAGGCAAT 2263

Db       4976 GTCCTTTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAGTCACAGGCAAT 5035

Qy       2264 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 2323

Db       5036 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 5095

Qy       2324 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 2383

Db       5096 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 5155

Qy       2384 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 2443

Db       5156 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAAGGAAGACACACTCGGTCTTCTT 5215

Qy       2444 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 2503

Db       5216 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 5275

Qy       2504 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 2563

Db       5276 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 5335

Qy       2564 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 2623

Db       5336 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 5395

Qy       2624 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGA 2682

Db       5396 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGA 5455

Qy       2683 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 2742

Db       5456 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 5515

Qy       2743 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 2802

Db       5516 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 5575

Qy       2803 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 2862

Db       5576 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 5635

Qy       2863 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 2922

Db       5636 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 5695

Qy       2923 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 2982

Db       5696 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 5755

Qy       2983 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 3042

Db       5756 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 5815

Qy       3043 TCG 3045

Db       5816 TCG 5818
```

RESULT 5
US-60-205-423-94
; Sequence 94, Application US/60205423
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000563
; CURRENT APPLICATION NUMBER: US/60/205,423
; CURRENT FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 494
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 94
; LENGTH: 2112
; TYPE: DNA
; ORGANISM: HUMAN
US-60-205-423-94

Query Match 45.4%; Score 1381.4; DB 76; Length 2112;
Best Local Similarity 99.6%; Pred. No. 5e-230;
Matches 1437; Conservative 0; Mismatches 1; Indels 5; Gaps 5;

```
Qy       1606 CAGATACTTCAGGCTTTCGAAGGAAGCT-AAAACATATAATAGGCCTGATATATAAGGTT 1664

Db          2 CAGATACTTCAGGCTTTCGAAGGAAGCTAAAAACATATAATAGGCCTGATATATAAGGTT 61

Qy       1665 CAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATA-GGAAGGAGACAAAACTC 1723

Db         62 CAGAGCAAAAGAGGGCACTAAAATAAATTTTTAAAGAAAATAGGGAAGGAGACAAAACTC 121

Qy       1724 AATACTACCTTGTCTTTTAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCA 1783

Db        122 AATACTACCTTGTCTTTTAATAACTGTCTTCCTCTTTCTAAAAGTTGTTGTATTTCCTCA 181

Qy       1784 ATACTTGCTTCATTTCTGGCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAA 1843

Db        182 ATACTTGCTTCATTTCT-GCATAAGGGTTTCCAGATGGCACTCTAAAACATTTGCCAGAA 240

Qy       1844 GGTGGCGAACATAAAACCTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGA 1903

Db        241 GGTGGCG-ACATAAAACCTCATTGCTTAGAACTGTCCCAGGTGCTGAACCCAGTTTCTGA 299
```

```
Qy       1904 GATTAAGAGAGGCTAGCCGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTT 1963
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GATTAAGAGAGGCTAGCCGGCTAGCGAACCGGGATTCCACCAAGTTTCCCCCAGAGGTTT 359

Qy       1964 GCAGGCTCTGGTAAGAAGTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACA 2023
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GCAGGCTCTGGTAAGAAGTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACA 419

Qy       2024 TATGCAAAATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGT 2083
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 TATGCAAAATATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGT 479

Qy       2084 CACTGAAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCA 2143
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CACTGAAATTACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCA 539

Qy       2144 CACTGAGTTTTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACAC 2203
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 CACTGAGTTTTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACAC 599

Qy       2204 GTCCTTTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAGTCACAGGCAAT 2263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GTCCTTTATAATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAGTCACAGGCAAT 659

Qy       2264 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 2323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 719

Qy       2324 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 2383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 779

Qy       2384 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 2443
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        780 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAAGGAAGACACACTCGGTCTTCTT 839

Qy       2444 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 2503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 899

Qy       2504 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 2563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 959

Qy       2564 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 2623
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 1019

Qy       2624 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGA 2682
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| |||
Db       1020 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGA 1079

Qy       2683 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 1139

Qy       2743 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 2802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 1199

Qy       2803 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 1259

Qy       2863 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 2922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 1319

Qy       2923 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 2982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 1379
```

```
Qy       2983 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 3042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 1439

Qy       3043 TCG 3045
             |||
Db       1440 TCG 1442
```

RESULT 6
PCT-US00-40519-1
; Sequence 1, Application PC/TUS0040519
; GENERAL INFORMATION:
; APPLICANT: Denten, R. Rex
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Kleim, Stepfanie
; APPLICANT: Stephens, J. Claiborne
; TITLE OF INVENTION: DRUG TARGET ISOGENES: POLYMORPHISMS IN THE
; TITLE OF INVENTION: 5-HYDROXYTRYPTAMINE RECEPTOR 1A GENE
; FILE REFERENCE: MWH-0006PCT HTR1A
; CURRENT APPLICATION NUMBER: PCT/US00/40519
; CURRENT FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: 60/147,711
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1204
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US00-40519-1

Query Match 24.6%; Score 747.8; DB 1; Length 1204;
Best Local Similarity 98.9%; Pred. No. 7.9e-120;
Matches 774; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

```
Qy       2264 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 2323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAGCGGAGGTACCGTTTTGTTGT 60

Qy       2324 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 2383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 120

Qy       2384 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 2443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 180

Qy       2444 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 2503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 240

Qy       2504 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 2563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 300

Qy       2564 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 2623
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTGGCAGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 360

Qy       2624 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGA 2682
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| |||
Db        361 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGA 420

Qy       2683 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 480

Qy       2743 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 2802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 540

Qy       2803 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 2862
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||||
Db        541 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCGCTTTCGAAGCGAC 599

Qy       2863 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 2922
                ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        600 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTAAAGGGAAGGCAGG 659

Qy       2923 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 2982
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 719

Qy       2983 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 3042
                ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        720 GAGGGAAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 779

Qy       3043 TCG 3045
                |||
Db        780 TCG 782


RESULT 7
US-10-049-407-1
; Sequence 1, Application US/10049407
; GENERAL INFORMATION:
;  APPLICANT: Denton, R. Rex
;  APPLICANT:  Kliem, Stefanie
;  APPLICANT:  Nandabalan, Krishnan
;  APPLICANT:  Stephens, J. Claiborne
;  TITLE OF INVENTION: DRUG TARGET ISOGENES: POLYMORPHISMS IN THE
;  TITLE OF INVENTION:  5-HYDROXYTRYPTAMINE RECEPTOR 1A GENE
;  FILE REFERENCE: MWH-0006US HTR1A
;  CURRENT APPLICATION NUMBER: US/10/049,407
;  CURRENT FILING DATE:  2002-02-06
;  PRIOR APPLICATION NUMBER: PCT/US00/40519
;  PRIOR FILING DATE: 2000-08-06
;  PRIOR APPLICATION NUMBER: 60/147,711
;  PRIOR FILING DATE: 1999-08-06
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1204
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-049-407-1

  Query Match             24.6%;  Score 747.8;  DB 43;  Length 1204;
  Best Local Similarity   98.9%;  Pred. No. 7.9e-120;
  Matches  774;  Conservative    0;  Mismatches    7;  Indels    2;  Gaps    2;

Qy       2264 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 2323
                ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAGCGGAGGTACCGTTTTGTTGT 60

Qy       2324 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 2383
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 120

Qy       2384 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 2443
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 180

Qy       2444 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 2503
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 240

Qy       2504 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 2563
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 300

Qy       2564 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 2623
                ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTGGCAGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 360

Qy       2624 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGA 2682
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        361 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGA 420

Qy       2683 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 2742
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 480

Qy       2743 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 2802
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 540

Qy       2803 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 2862
                ||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||
Db        541 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCGCTTTCGAAGCGAC 599

Qy       2863 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 2922
                ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        600 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTAAAGGGAAGGCAGG 659

Qy       2923 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 2982
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 719

Qy       2983 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 3042
                ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        720 GAGGGAAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 779

Qy       3043 TCG 3045
                |||
Db        780 TCG 782


RESULT 8
PCT-US00-40519-3
; Sequence 3, Application PC/TUS0040519
; GENERAL INFORMATION:
;  APPLICANT:  Denten, R. Rex
;  APPLICANT:  Nandabalan, Krishnan
;  APPLICANT:  Kleim, Stepfanie
;  APPLICANT:  Stephens, J. Claiborne
;  TITLE OF INVENTION:  DRUG TARGET ISOGENES: POLYMORPHISMS IN THE
;  TITLE OF INVENTION:  5-HYDROXYTRYPTAMINE RECEPTOR 1A GENE
;  FILE REFERENCE:  MWH-0006PCT HTR1A
;  CURRENT APPLICATION NUMBER: PCT/US00/40519
;  CURRENT FILING DATE:  2000-08-01
;  PRIOR APPLICATION NUMBER:  60/147,711
;  PRIOR FILING DATE:  1999-08-06
;  NUMBER OF SEQ ID NOS:  53
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 2722
;   TYPE: DNA
;   ORGANISM: Homo sapiens
PCT-US00-40519-3

  Query Match             24.6%;  Score 747.8;  DB 1;  Length 2722;
  Best Local Similarity   98.9%;  Pred. No. 9.7e-120;
  Matches  774;  Conservative    0;  Mismatches    7;  Indels    2;  Gaps    2;

Qy       2264 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 2323
                ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAGCGGAGGTACCGTTTTGTTGT 60

Qy       2324 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 2383
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 120

Qy       2384 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 2443
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 180

Qy       2444 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 2503
```

```
Db        181 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 240

Qy       2504 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 2563

Db        241 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 300

Qy       2564 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 2623

Db        301 TTGGCAGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 360

Qy       2624 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGA 2682

Db        361 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGA 420

Qy       2683 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 2742

Db        421 CCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 480

Qy       2743 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 2802

Db        481 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 540

Qy       2803 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 2862

Db        541 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCGCTTTCGAAGCGAC 599

Qy       2863 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 2922

Db        600 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTAAAGGGAAGGCAGG 659

Qy       2923 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 2982

Db        660 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 719

Qy       2983 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 3042

Db        720 GAGGGAAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 779

Qy       3043 TCG 3045

Db        780 TCG 782
```

RESULT 9
US-10-049-407-3
; Sequence 3, Application US/10049407
; GENERAL INFORMATION:
; APPLICANT: Denton, R. Rex
; APPLICANT: Kliem, Stefanie
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Stephens, J. Claiborne
; TITLE OF INVENTION: DRUG TARGET ISOGENES: POLYMORPHISMS IN THE
; TITLE OF INVENTION: 5-HYDROXYTRYPTAMINE RECEPTOR 1A GENE
; FILE REFERENCE: MWH-0006US HTR1A
; CURRENT APPLICATION NUMBER: US/10/049,407
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: PCT/US00/40519
; PRIOR FILING DATE: 2000-08-06
; PRIOR APPLICATION NUMBER: 60/147,711
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2722
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-049-407-3

Query Match 24.6%; Score 747.8; DB 43; Length 2722;
Best Local Similarity 98.9%; Pred. No. 9.7e-120;
Matches 774; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

```
Qy       2264 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGT 2323

Db          1 ATTCTCCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAGCGGAGGTACCGTTTTGTTGT 60

Qy       2324 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 2383

Db         61 TGTTGTCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGG 120

Qy       2384 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 2443

Db        121 AGTGCAATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTT 180

Qy       2444 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 2503

Db        181 CCATCAATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTC 240

Qy       2504 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 2563

Db        241 CCTATGCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAAT 300

Qy       2564 TTGGCAGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 2623

Db        301 TTGGCAGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCC 360

Qy       2624 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGA 2682

Db        361 CAGTTTTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGA 420

Qy       2683 TCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 2742

Db        421 CCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGA 480

Qy       2743 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 2802

Db        481 CGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTG 540

Qy       2803 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGAC 2862

Db        541 GGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCGCTTTCGAAGCGAC 599

Qy       2863 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGG 2922

Db        600 TCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTAAAGGGAAGGCAGG 659

Qy       2923 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 2982

Db        660 TGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGA 719

Qy       2983 GAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 3042

Db        720 GAGGGAAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTC 779

Qy       3043 TCG 3045

Db        780 TCG 782
```

RESULT 10
US-09-541-946-2477
; Sequence 2477, Application US/09541946
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Cargill, Michele
; APPLICANT: Altshuler, David M.
; APPLICANT: Ireland, James S.
; APPLICANT: Sklar, Pamela
; APPLICANT: Patil, Nila
; APPLICANT: Lipshutz, Robert J.
; APPLICANT: Daley, George Q.
; TITLE OF INVENTION: CHARACTERIZATION OF SINGLE NUCLEOTIDE
; TITLE OF INVENTION: POLYMORPHISMS IN CODING REGIONS OF HUMAN GENES
; FILE REFERENCE: 2825.1017-003
; CURRENT APPLICATION NUMBER: US/09/541,946

```
;  CURRENT FILING DATE:  2000-03-31
;  PRIOR APPLICATION NUMBER: US 60/127,248
;  PRIOR FILING DATE: 1999-03-31
;  NUMBER OF SEQ ID NOS: 2889
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2477
;   LENGTH: 839
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-541-946-2477

  Query Match             24.4%;  Score 742.4;  DB 23;  Length 839;
  Best Local Similarity   98.7%;  Pred. No. 6.3e-119;
  Matches  768;  Conservative    1;  Mismatches    7;  Indels    2;  Gaps    2;

Qy       2269 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGTTGTTG 2328
             ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db             1 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAGCGGAGGTACCGTTTTGTTGTTGTTG 60

Qy       2329 TCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGC 2388
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db            61 TCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTYGCCCAGGCTGGAGTGC 120

Qy       2389 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATC 2448
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATC 180

Qy       2449 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 2508
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 240

Qy       2509 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 2568
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           241 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 300

Qy       2569 AGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 2628
             |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db           301 AGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 360

Qy       2629 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGATCTGG 2687
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||||||
Db           361 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGACCTGG 420

Qy       2688 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 2747
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           421 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 480

Qy       2748 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 2807
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           481 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 540

Qy       2808 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGACTCACA 2867
             |||||||||||||||||||||||||||||||| || ||||||||||||||||||||||||
Db           541 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCGCTTTCGAAGCGACTCACA 599

Qy       2868 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGGTGGGG 2927
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db           600 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTAAAGGGAAGGCAGGTGGGG 659

Qy       2928 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 2987
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           660 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 719

Qy       2988 AAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 3045
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db           720 AAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 777


RESULT 11
US-09-541-946-2478
; Sequence 2478, Application US/09541946
; GENERAL INFORMATION:
;  APPLICANT: Lander, Eric S.
;  APPLICANT:  Cargill, Michele
;  APPLICANT:  Altshuler, David M.
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Sklar, Pamela
;  APPLICANT:  Patil, Nila
;  APPLICANT:  Lipshutz, Robert J.
;  APPLICANT:  Daley, George Q.
;  TITLE OF INVENTION: CHARACTERIZATION OF SINGLE NUCLEOTIDE
;  TITLE OF INVENTION:  POLYMORPHISMS IN CODING REGIONS OF HUMAN GENES
;  FILE REFERENCE: 2825.1017-003
;  CURRENT APPLICATION NUMBER: US/09/541,946
;  CURRENT FILING DATE:  2000-03-31
;  PRIOR APPLICATION NUMBER: US 60/127,248
;  PRIOR FILING DATE: 1999-03-31
;  NUMBER OF SEQ ID NOS: 2889
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2478
;   LENGTH: 839
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-541-946-2478

  Query Match             24.4%;  Score 742.4;  DB 23;  Length 839;
  Best Local Similarity   98.7%;  Pred. No. 6.3e-119;
  Matches  768;  Conservative    1;  Mismatches    7;  Indels    2;  Gaps    2;

Qy       2269 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGTTGTTG 2328
             ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db             1 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAGCGGAGGTACCGTTTTGTTGTTGTTG 60

Qy       2329 TCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGC 2388
             ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db            61 TCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGYTGGAGTGC 120

Qy       2389 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATC 2448
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATC 180

Qy       2449 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 2508
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 240

Qy       2509 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 2568
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           241 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 300

Qy       2569 AGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 2628
             |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db           301 AGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 360

Qy       2629 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGATCTGG 2687
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||||||
Db           361 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGACCTGG 420

Qy       2688 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 2747
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           421 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 480

Qy       2748 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 2807
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           481 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 540

Qy       2808 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGACTCACA 2867
             |||||||||||||||||||||||||||||||| || ||||||||||||||||||||||||
Db           541 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCGCTTTCGAAGCGACTCACA 599

Qy       2868 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGGTGGGG 2927
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db           600 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTAAAGGGAAGGCAGGTGGGG 659

Qy       2928 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 2987
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           660 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 719
```

```
Qy       2988 AAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 3045
             ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        720 AAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 777
```

RESULT 12
US-09-541-946-2479
; Sequence 2479, Application US/09541946
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Cargill, Michele
; APPLICANT: Altshuler, David M.
; APPLICANT: Ireland, James S.
; APPLICANT: Sklar, Pamela
; APPLICANT: Patil, Nila
; APPLICANT: Lipshutz, Robert J.
; APPLICANT: Daley, George Q.
; TITLE OF INVENTION: CHARACTERIZATION OF SINGLE NUCLEOTIDE
; TITLE OF INVENTION: POLYMORPHISMS IN CODING REGIONS OF HUMAN GENES
; FILE REFERENCE: 2825.1017-003
; CURRENT APPLICATION NUMBER: US/09/541,946
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 60/127,248
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 2889
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2479
; LENGTH: 839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-541-946-2479

Query Match 24.4%; Score 742.4; DB 23; Length 839;
Best Local Similarity 98.7%; Pred. No. 6.3e-119;
Matches 768; Conservative 1; Mismatches 7; Indels 2; Gaps 2;

```
Qy       2269 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGTTGTTG 2328
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db          1 CCCTGAGGGAGTAAGGCTGGACTGTTAGATGATAGCGGAGGTACCGTTTTGTTGTTGTTG 60

Qy       2329 TCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGC 2388
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCGTCGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGC 120

Qy       2389 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATC 2448
             ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        121 AATGGCGCGAGAACGGAGGTAGCTTTTTAAAAASGAAGACACACTCGGTCTTCTTCCATC 180

Qy       2449 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 2508
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATTAGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTAT 240

Qy       2509 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 2568
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCTTCCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGC 300

Qy       2569 AGATAATATGAGGCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 2628
             |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGATAATATGACCCAAGGAGTAGTTGGAATTCCCTCCCCCAAGTTTTTCCAACCCCAGTT 360

Qy       2629 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCA-GGATCTGG 2687
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||||
Db        361 TTGCTGGGTTGGAGGCGGAGTTTATTTGTTACAACCTTGGTCTGACCGGCAGGGACCTGG 420

Qy       2688 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 2747
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAAGAGACTCGAATGCAAAGACGCTG 480

Qy       2748 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 2807
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAGAGGCACTCCTCGGGGTTGGGGAA 540
```

```
Qy       2808 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAGCCTGGCTTTCGAAGCGACTCACA 2867
             |||||||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db        541 GTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG-CTCGCTTTCGAAGCGACTCACA 599

Qy       2868 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTGAAAGGGAAGGCAGGTGGGG 2927
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAGACTTAAAGGGAAGGCAGGTGGGG 659

Qy       2928 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 2987
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 AGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGAGGAGGAGAGAGGGGGAGAGAGGG 719

Qy       2988 AAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 3045
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        720 AAGGAAGGAAATAGCGAGAGGAGGGTCACAGAGTGACCGTGGAGGATGGGGCTTCTCG 777
```

RESULT 13
US-09-634-306B-196186/c
; Sequence 196186, Application US/09634306B
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/09/634,306B
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 196186
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Human
US-09-634-306B-196186

Query Match 14.1%; Score 428.2; DB 26; Length 559;
Best Local Similarity 92.2%; Pred. No. 2.7e-64;
Matches 517; Conservative 2; Mismatches 30; Indels 12; Gaps 6;

```
Qy        599 ATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTTGA 658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 ATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTTGA 500

Qy        659 CACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTCCT 718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 CACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTCCT 440

Qy        719 TTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTCAG 778
              |||||||||| ||||||||||||||||:|||||||||||||||||||||||||||||||
Db        439 TTATTTTACTTGGTTTATAGTCTCAACTATAMTTTCAAAGTTTAATTTTAGATAATTCAG 380

Qy        779 CCTTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTTAACTGTAAAT 838
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        379 -CTTTTAAATATTTTCCCATTATAATTTTTG-GACCTCTAACTCTATTTTAACTGTAAAT 322

Qy        839 ATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTT-A 897
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        321 ACAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTTAA 262
```

```
Qy        898 AAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTGCAT 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        261 AAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTTCAA 202

Qy        958 GCC----ATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTG-AATGATTGAGTGG 1012
                  |    |      || |  ||| |     ||||||||||||||| |||||||||||||
Db        201 TGCCATGAATCATCACAAATGGCAGGCTCATTGTGGTGGCATGCTGAAATGATTGAGTGG 142

Qy       1013 GACTGTGCC----AGCTGAACTATAAAAAAAAAAAAACAAACAAAACCTTATCCAAACACA 1068
              ||       |      ||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GAACTGTGCCCAGCACTGAACTATAAAAAAAAAAAAACAAACAAAACCTTATCCAAACACA 82

Qy       1069 CTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTGCT 1128
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db         81 CTGTCCTGTATTGTAATGCATTGGCCCAASTGGATTCTTTTTGATGCTTTGGTGATTGCT 22

Qy       1129 CTTTTGTTTGGGCTTGGAGAA 1149
              |||||||||||||||||||||
Db         21 CTTTTGTTTGGGCTTGGAGAA 1
```

RESULT 14
US-09-634-306B-196187/c
; Sequence 196187, Application US/09634306B
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/09/634,306B
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 196187
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Human
US-09-634-306B-196187

Query Match 14.1%; Score 428.2; DB 26; Length 559;
Best Local Similarity 92.2%; Pred. No. 2.7e-64;
Matches 517; Conservative 2; Mismatches 30; Indels 12; Gaps 6;

```
Qy        599 ATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTTGA 658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 ATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTTGA 500

Qy        659 CACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTCCT 718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 CACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTCCT 440

Qy        719 TTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTCAG 778
              |||||||||||| ||||||||||||||||||:||||||||||||||||||||||||||||
Db        439 TTATTTTACTTGGTTTATAGTCTCAACTATAMTTTCAAAGTTTAATTTTAGATAATTCAG 380

Qy        779 CCTTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTTAACTGTAAAT 838
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        379 -CTTTTAAATATTTTCCCATTATAATTTTTG-GACCTCTAACTCTATTTTAACTGTAAAT 322
```

```
Qy        839 ATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTT-A 897
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        321 ACAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTTAA 262

Qy        898 AAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTGCAT 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        261 AAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTTCAA 202

Qy        958 GCC----ATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTG-AATGATTGAGTGG 1012
                  |    |      || |  ||| |     ||||||||||||||| |||||||||||||
Db        201 TGCCATGAATCATCACAAATGGCAGGCTCATTGTGGTGGCATGCTGAAATGATTGAGTGG 142

Qy       1013 GACTGTGCC----AGCTGAACTATAAAAAAAAAAAAACAAACAAAACCTTATCCAAACACA 1068
              ||       |      ||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GAACTGTGCCCAGCACTGAACTATAAAAAAAAAAAAACAAACAAAACCTTATCCAAACACA 82

Qy       1069 CTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTGCT 1128
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db         81 CTGTCCTGTATTGTAATGCATTGGCCCAASTGGATTCTTTTTGATGCTTTGGTGATTGCT 22

Qy       1129 CTTTTGTTTGGGCTTGGAGAA 1149
              |||||||||||||||||||||
Db         21 CTTTTGTTTGGGCTTGGAGAA 1
```

RESULT 15
US-10-027-632-196186/c
; Sequence 196186, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 196186
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-196186

Query Match 14.1%; Score 428.2; DB 43; Length 559;
Best Local Similarity 92.2%; Pred. No. 2.7e-64;
Matches 517; Conservative 2; Mismatches 30; Indels 12; Gaps 6;

```
Qy        599 ATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTTGA 658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 ATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTTGA 500

Qy        659 CACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTCCT 718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 CACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTCCT 440

Qy        719 TTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTCAG 778
              |||||||||||| ||||||||||||||||||:||||||||||||||||||||||||||||
Db        439 TTATTTTACTTGGTTTATAGTCTCAACTATAMTTTCAAAGTTTAATTTTAGATAATTCAG 380
```

```
Qy        779 CCTTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTTAACTGTAAAT 838
              ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        379 -CTTTTAAATATTTTCCCATTATAATTTTTG-GACCTCTAACTCTATTTTAACTGTAAAT 322

Qy        839 ATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTT-A 897
              | |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        321 ACAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTTAA 262

Qy        898 AAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTGCAT 957
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        261 AAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTTCAA 202

Qy        958 GCC----ATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTG-AATGATTGAGTGG 1012
                 |    |       || |  ||| |     |||||||||||||| ||||||||||||
Db        201 TGCCATGAATCATCACAAATGGCAGGCTCATTGTGGTGGCATGCTGAAATGATTGAGTGG 142

Qy       1013 GACTGTGCC----AGCTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACACA 1068
              ||      |     ||||||||||||||||||||||||||||||||||||||||||||
Db        141 GAACTGTGCCCAGCACTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACACA 82

Qy       1069 CTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTGCT 1128
              ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db         81 CTGTCCTGTATTGTAATGCATTGGCCCAASTGGATTCTTTTTGATGCTTTGGTGATTGCT 22

Qy       1129 CTTTTGTTTGGGCTTGGAGAA 1149
              |||||||||||||||||||||
Db         21 CTTTTGTTTGGGCTTGGAGAA 1
```

Search completed: April 10, 2004, 17:18:56
Job time : 8790 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 08:29:12 ; Search time 6908 Seconds
(without alignments)
13163.049 Million cell updates/sec

Title: US-09-430-412A-1
Perfect score: 3045
Sequence: 1 atcatcaataatatccgtta.....gtggaggatggggcttctcg 3045

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	594	19.5	635	29	AG174550	AG174550 Pan trogl
	2	420	13.8	515	28	AQ885016	AQ885016 HS_5518_A
	3	111.2	3.7	1101	29	CNS00EVL	AL069706 Drosophil
c	4	108.6	3.6	475	28	AQ304345	AQ304345 HS_3176_B
c	5	107.2	3.5	1101	29	CNS00EVL	AL069706 Drosophil
c	6	103.6	3.4	1201	9	AL565455	AL565455 AL565455
c	7	102	3.3	1056	13	BX415058	BX415058 BX415058
c	8	101.4	3.3	1201	9	AL536104	AL536104 AL536104
c	9	101	3.3	1165	13	BX338369	BX338369 BX338369
c	10	100.8	3.3	714	28	AZ849216	AZ849216 2M0150L03
c	11	100.4	3.3	1098	13	BX377526	BX377526 BX377526
	12	99.6	3.3	1200	13	BX415878	BX415878 BX415878
c	13	99.2	3.3	1101	29	CNS00LT2	AL078714 Drosophil
	14	98.8	3.2	1272	28	CC264939	CC264939 CH261-19L
c	15	98.4	3.2	1101	29	CNS0021J	AL061936 Drosophil
c	16	96	3.2	997	29	CNS005TE	AL060767 Drosophil
c	17	94.6	3.1	994	29	CNS04NOJ	AL298972 Tetraodon
	18	94.2	3.1	1201	13	BX335216	BX335216 BX335216
c	19	93.8	3.1	1201	13	BX458623	BX458623 BX458623
c	20	93.8	3.1	1295	29	CG757539	CG757539 P052-4-E0
c	21	92.4	3.0	1101	29	CNS00EO7	AL069440 Drosophil
	22	92.4	3.0	1364	29	CG757970	CG757970 P053-2-C0
c	23	92.2	3.0	1101	29	CNS0039G	AL063921 Drosophil
c	24	92	3.0	1200	13	BX437758	BX437758 BX437758
c	25	91.8	3.0	1201	13	BX448679	BX448679 BX448679
	26	91.4	3.0	1200	13	BX437758	BX437758 BX437758
c	27	91.2	3.0	872	14	CK022958	CK022958 AGENCOURT
c	28	91.2	3.0	1201	13	BX443774	BX443774 BX443774
	29	91	3.0	1045	13	BX456814	BX456814 BX456814
c	30	90.6	3.0	1201	9	AL547503	AL547503 AL547503
c	31	90.6	3.0	1201	13	BX357882	BX357882 BX357882
c	32	90	3.0	987	29	CNS00418	AL066537 Drosophil
c	33	89.8	2.9	1052	13	BU560001	BU560001 AGENCOURT
c	34	89.2	2.9	1200	13	BX415878	BX415878 BX415878
	35	88.8	2.9	975	28	BH179465	BH179465 014_P_10-
	36	88.8	2.9	1201	9	AL536104	AL536104 AL536104
	37	88.8	2.9	1391	29	CG754863	CG754863 P050-2-G0
	38	88.8	2.9	1392	29	CG757503	CG757503 P052-4-C0
c	39	88.6	2.9	1200	13	BX436510	BX436510 BX436510
	40	88.2	2.9	1101	29	CNS017RP	AL108415 Drosophil
	41	88	2.9	1592	29	CG750135	CG750135 P044-3-D0
	42	87.8	2.9	1201	13	BX443774	BX443774 BX443774
c	43	87.8	2.9	1201	13	BX461824	BX461824 BX461824
	44	87.8	2.9	1350	29	CG744271	CG744271 P036-4-C0
c	45	87.4	2.9	1140	13	BX415974	BX415974 BX415974

ALIGNMENTS

RESULT 1
AG174550/c
LOCUS AG174550 635 bp DNA linear GSS 09-JAN-2002
DEFINITION Pan troglodytes DNA, clone: RP43-044N13.TJ, genomic survey sequence.
ACCESSION AG174550
VERSION AG174550.1 GI:16704230
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library RPCI-43
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 635)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library RPCI-43 This BAC

```
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                 Sequencing: TJ
             LIBRARY
                 Vector    : pBACe3.6
                 R.Site 1  : EcoRI
                 R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .635
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-044N13.TJ"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
ORIGIN

  Query Match            19.5%;  Score 594;  DB 29;  Length 635;
  Best Local Similarity   97.5%;  Pred. No. 1.1e-91;
  Matches  614;  Conservative    0;  Mismatches   15;  Indels    1;  Gaps    1;

Qy       1973 GGTAAGAAGTGCAAAAGGCCATGTGAAATGCCAGGCTTCACTTAGAACACATATGCAAAA 2032
             |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        635 GGTAAGAAGTGCAAAAGGCCATGTAAAATGCCAGGCTTCACTTAGAACACATATGCAAAA 576

Qy       2033 TATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGTCACTGAAAT 2092
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        575 TATTTCCATCCCTGAATTTACTAGCCACAAAGCTATGGGAAGTGGCAGTGTCACTGAAAT 516

Qy       2093 TACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCACACTGAGTT 2152
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        515 TACAAGTGTAGTAGTGATGGAAAAGTGTGTGTGTGTTTAGAATATATATCACACTGAGTT 456

Qy       2153 TTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACACGTCCTTTAT 2212
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 TTGTTCTTCATTTCGAGATGCAGTTGTTTACCTCTCCTTGTCCTTTGACACGTCCTTTAT 396

Qy       2213 AATTTCGTTCTCTCCCGGTTCCCCAACGTTAAAAAAAAAGTCACAGGCAATATTCTCCCT 2272
             |||||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||
Db        395 AATTTCGTTCTCTCCCAGTTCCCCAGGGTTAAAAAAAAAGTCACAGGCAATATTCTCCCT 336

Qy       2273 GAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGTTGTTGTCGT 2332
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 GAGGGAGTAAGGCTGGACTGTTAGATGATAACGGAGGTACCGTTTTGTTGTTGTTGTCGT 276

Qy       2333 CGTTGTTCGTTTGTTTTTGGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAATG 2392
             |||||||||||||||||| ||||| |||||||||||||||||||||||||||||||||| |
Db        275 CGTTGTTCGTTTGTTTTTAGAGACAGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAACG 216

Qy       2393 GCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATCAATT 2452
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 GCGCGAGAACGGAGGTAGCTTTTTAAAAACGAAGACACACTCGGTCTTCTTCCATCAATT 156

Qy       2453 AGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATTCAGGCTCCCTATGCTT 2512
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 AGCAATAATTGGGAGACTGACCCAGGACTGTTCACCTTCCCATGCAGGCTCCCTATGCTT 96

Qy       2513 CCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGCAGAT 2572
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 CCTTTTCTCATCTCCTATTGCCACTCTGGGATGCTGACACGATTTAAGAATTTGGCAGAT 36

Qy       2573 AATATGAGGCAAGGAGTAGTTGGAATTCCC 2602
             ||||| |||||||||| || | | | | |
Db         35 AATAT-AGGCAAGGAGATGTGGAATTCCGC 7


RESULT 2
AQ885016
LOCUS       AQ885016                 515 bp    DNA     linear   GSS 09-NOV-1999
DEFINITION  HS_5518_A2_C02_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=9286 Col=4 Row=E, genomic survey sequence.
ACCESSION   AQ885016
VERSION     AQ885016.1  GI:6316483
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 515)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 9286  row: E  column: 4
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 515.
FEATURES             Location/Qualifiers
     source          1. .515
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=9286 Col=4 Row=E"
                     /sex="male"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
ORIGIN

  Query Match            13.8%;  Score 420;  DB 28;  Length 515;
  Best Local Similarity   96.2%;  Pred. No. 6.7e-62;
  Matches  459;  Conservative    0;  Mismatches   15;  Indels    3;  Gaps    3;

Qy       1163 AATTCAGAGCTCAGATTTGAACACAATATTAAGATTATTGCAATCTGTAGTGAATCTGTT 1222
             ||  || |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         32 AANNCANAGCTCAGATNTGAACACAATATTAAGATTATTGCAATCTGTAGTGAATCTGTT 91

Qy       1223 CATGTTATCCAGTGTCAACTGCTTTTGAGATTGCATTCCTTTCACCTCAGGCATGCAATC 1282
             |||||||||||||||||||||||| |||||||||||||||||||||||| ||||||||||
Db         92 CATGTTATCCAGTGTCAACTGCNTTTGAGATTGCATTCCTTTCACCTCAAGCATGCAATC 151

Qy       1283 AGGATGTATAAGTGAAATGTTGTGTGGTATGTTTACTGTAGTTGCTTAGAAGT-CCATTC 1341
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        152 AGGATGTATAAGTGAAATGTTGTGTGGTATGTTTACTGTAGTTGCTTAGAAGTCCCATTC 211

Qy       1342 TTTACCAATGCTCAAATGTGATTAAATTTG-TTTTCTTGTTAAAGGAAACAGCTTAGAAC 1400
             |||||||||||||||||||||||||||||| |||||||||||||||||||||| |||
Db        212 TTTACCAATGCTCAAATGTGATTAAATTTGTTTTTCTTGTTAAAGGAAACAGCTTATAAC 271

Qy       1401 AAACCCTTGTAAGTATCTTTA-TTTCAGTGATTTAACATTTCCAAATGTTAAATCATTTG 1459
             |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        272 AAACCCTTGTAAGTATCTTTATTTTCAGTGATTTAACATTTCCAAATGTTAAATCATTTG 331

Qy       1460 GAAAATGCAATACTATTCGTTTCTCCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTC 1519
```

```
Db        332 GAAAATGCAATACTATTCTTTTCTCCAACAAAAGGTAAATTTATGTCAGTTCCAAAGTTC 391

Qy       1520 AGGTTATGACAGCACAAAACCAACACAGGTGAAAGTGTTAGCCTAGCTTTATTAAAATGG 1579

Db        392 ACGTTATGACAGCACAAAACCAACACAGGTGAGAGTGTTAGCCTAGCTTTATTACAATGG 451

Qy       1580 CATTCCCAGTTAGAACTTGTGAATGACAGATACTTCAGGCTTTCGAAGGAAGCTAAA 1636

Db        452 CATTCCCACTTAGAACTTGTGAATGACAGATACTTCACGCTTTCTAAAGAAGCTAAA 508


RESULT 3
CNS00EVL
LOCUS       CNS00EVL                1101 bp    DNA     linear   GSS 04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC:
            BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL069706
VERSION     AL069706.1  GI:4949849
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR29B23"
                     /clone_lib="RPCI-98"
                     /note="end : T7"
ORIGIN

  Query Match             3.7%;  Score 111.2;  DB 29;  Length 1101;
  Best Local Similarity   35.7%;  Pred. No. 3.3e-09;
  Matches  232;  Conservative  127;  Mismatches  286;  Indels    5;  Gaps    2;

Qy        126 TTCTGTATCTTTAAGAGTTAAACATAGAAGAATTGGCTAAGTGAAAATGAATGAAACGCA 185

Db        454 YTCTCAHTWTMMMMMWWAATWTWWAAAWAAAWTTATWAATWAAAAAWAWWWWATTTTTW 513

Qy        186 ATATCATTCTGCATATATCATTTATTATATATCACAGTATTATTAGTTTTAAAAGTTAAA 245

Db        514 WWWWTWATTWTTTWAWWTWTAWTAAAAAAAAAWATAATTTAAAWWAATAWATTAAWAATT 573

Qy        246 CATAAATATCTATTATGYCATTGSACGAYTAGGYSAACCTARTCRGTGCTGCGAATACTT 305

Db        574 TAWAAWWTATATTAATWTATAAATWTWATTAATATAAAAAAATATTTTTTWATAAAATTT 633

Qy        306 TCGATACTTCTGTTTCCCTCCTAGTATTCATAAGTGTGCCTTTGAAAACGTTTTAAATTG 365

Db        634 TTAATAATTTAATT--AWTTATTAAATAAWTWATTWWWTAATTAAATAATTTWAAATAWA 691

Qy        366 TAAGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACTTGAATTACT 425

Db        692 AAAAAAAAAAAAAAAWATWAAWAATWATAWATAAWTTAAAAWAATAAAAWAAWAATWAWWA 751

Qy        426 TTGGATTTTGAAAAACTTTGATAAATTCTACAT---CATAGCATATTGAAGCAAGAATAA 482

Db        752 TAATATWWATATATAWTTWTAWWWATWWAWWWTATAWATAWAATAWAAWAWAWATAAATA 811

Qy        483 CAAATGCTATACCTCAGGAATATTAATTCCAGATTTTACAGCATTTTAACTTTCTTGATG 542

Db        812 WATAWATWAAAWAWAWATAWWATWATATAWWAATAWAWAAAAAATWTAATATWAATWATA 871

Qy        543 AGAAAAAATAAATTTGTCAGTTATTAAACTATTTGGATCCAACAGATGAAAGCAGAATTC 602

Db        872 AWAAAAAATAWAWTTWWTWTTTTWWAWWATATAAAWAWATAWAAAWAAAAAAAAAAAAAT 931

Qy        603 TAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTTGACACA 662

Db        932 AAWAWWWTWATATTWTTATTAAAWTWTATWWATTWATTWWAWTWTATAWTTWTATWTATA 991

Qy        663 ATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTCCTTTAT 722

Db        992 TATWWTAWTAWATATATTTATTAAWWTATATTTTAWAAAWTAATATATWATAWWWTAWAT 1051

Qy        723 TTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATA 772

Db       1052 ATAWAWWAATTAWTTATATATWTAAWATAWWAAAAAWAAAWAWATAWWAA 1101


RESULT 4
AQ304345/c
LOCUS       AQ304345                475 bp    DNA     linear   GSS 16-DEC-1998
DEFINITION  HS_3176_B2_H11_T7 CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3176 Col=22 Row=P, genomic survey
            sequence.
ACCESSION   AQ304345
VERSION     AQ304345.1  GI:4020182
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 475)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3176  row: P  column: 22
            Class: BAC ends
            High quality sequence stop: 475.
FEATURES             Location/Qualifiers
     source          1. .475
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3176 Col=22 Row=P"
                     /sex="male"
```

/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

ORIGIN

Query Match 3.6%; Score 108.6; DB 28; Length 475;
Best Local Similarity 70.7%; Pred. No. 1.2e-08;
Matches 186; Conservative 0; Mismatches 68; Indels 9; Gaps 3;

```
Qy        886 TCACAAATCTTAAAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCT 945
              |||     | | ||||  || | ||||||     |     |||  ||  ||| |||   | |
Db        284 TCAATTTTGTGAAAAAGTTTTTACAGAGTGGTGAGCAGCCATACCCGGGTANACTAGTTT 225

Qy        946 CTTTCTTTGCATGCC----ATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTGAA 1001
              | || ||| ||  ||     ||  |   || |   || |      |||| ||||||||| ||
Db        224 CATTTTTTCCAACCCCATGATTCTTCCCAAATGCCAGGCTCATTGTGTTGGCATGCTAAA 165

Qy       1002 TG-ATTGAGTGGGACTGTGCC----AGCTGAACTATAAAAAAAAAAAACAAACAAAACCT 1056
               | |||||||||  ||||  |       ||||||||||||||||||  ||||||||||||||
Db        164 AGAATTGAGTGGAACTGTTGCCCAGCACTGAACTATAAAAAAAAAGAACAAACAAAACCT 105

Qy       1057 TATCCAAACACACTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCT 1116
              ||||||||||||||||||||||||||||||||||||||||||||||||||| | ||| | ||
Db        104 TATCCAAACACACTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTNTGTGGTNCT 45

Qy       1117 TTGGTGATTGCTCTTTTGTTTGG 1139
                 ||||| ||||||  ||||  |
Db          44 CGTGTGATAGCTCTTGTGTTGNG 22
```

RESULT 5
CNS00EVL/c
LOCUS CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC: BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL069706
VERSION AL069706.1 GI:4949849
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES Location/Qualifiers
source 1. .1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR29B23"
/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN

Query Match 3.5%; Score 107.2; DB 29; Length 1101;
Best Local Similarity 36.3%; Pred. No. 1.6e-08;
Matches 213; Conservative 108; Mismatches 261; Indels 4; Gaps 1;

```
Qy        331 ATTCATAAGTGTGCCTTTGAAAACGTTTTAAATTGTAAGAAATAAAATGTTTGATATATT 390
              ||| :|:  |  |   ::: | :|  | |||:  || ||| | |||::  |  |  ||||||:
Db       1061 ATTWWTWTATATWTAWWWTATWATATATTAWTTTWTAAAATATAWWTTAATAAATATATW 1002

Qy        391 ATGTATATTATTACTAAGAAA----AAACTTGAATTACTTTGGATTTTGAAAAACTTTGA 446
                ||:: |||   :|  :||    |:| |: |||:| |::  |: :|  || | :   |
Db       1001 TAWTAWWATATATAWATAWAAWTATAWAWTWWAATWAATWWATAWAWTTTAATAAWAATA 942

Qy        447 TAAATTCTACATCATAGCATATTGAAGCAAGAATAACAAATGCTATACCTCAGGAATATT 506
              |:|::: :  ||  |    | || :   :   :|:   |    : ::    :  :| :|:
Db        941 TWAWWWTWTTATTTTTTTTTTTTTWTTTWTATWTWTTTATATWWTWWAAAAWAWWAAWTW 882

Qy        507 AATTCCAGATTTTACAGCATTTTAACTTTCTTGATGAGAAAAAATAAATTTGTCAGTTAT 566
              |||    :|| |:    || |||: ||| ||  : |  :: |   :||::   : :|:|
Db        881 TATTTTTTWTTATWATTWATATTAWATTTTTTWTWTATTWWTATATWATWWTATWTWTWT 822

Qy        567 TAAACTATTTGGATCCAACAGATGAAAGCAGAATTCTAACTAACATATTTATTGATTTAT 626
              | :| :  |:   |  : : :| : |  : | :| ||: : :  ||:::||:  :| |||
Db        821 TTWATWTATWTATTTATWTWTWTTWTATTWTATWTATAWWWTWWATWWWTAWAAWTATAT 762

Qy        627 TTGTGATTTACATATTTACATGTGTTGTTTGACACAATTCTTAATTATGTTCTTGATATG 686
              | :   ||| :: : |   | | || ||:     |:|| |: ||:|| :| :  :| | |
Db        761 ATWWATATTATWWTWATTWTTWTTTTATTWTTTAAWTTATWTATWATTWTTWATWTTTT 702

Qy        687 CATATATTTGCTTCTTAAATTTTAAGTTTCCTTTATTTTACTTTGTTTATAGTCTCAACT 746
              | | |||  :| || :|  |||  |     :::  |:| ||  || ||| :
Db        701 TTTTTTTTTTTWTATTTWAAATTATTTAATTAWWWAATWAWTTATTTAATAAWTAATTAA 642

Qy        747 ATAATTTCAAAGTTTAATTTTAGATAATTCAGCCTTTTAAATATTTTCCCATTATAATTT 806
              || |||  ||| |||| :   | ||| ||     | |||| :|: ||   |:  |||| |
Db        641 ATTATTAAAAATTTTATWAAAAAATATTTTTTTATATTAATWAWATTTATAWATTAATAT 582

Qy        807 TTGTGACCTCTAACTCTATTTTAACTGTAAATATAGTTCTGTATTTGTGAAGAGACTTTA 866
              : | :    | : |   :| | : | |||||    || | | ||  | :| : :   :|
Db        581 AWWTTWTAAATTWTTAATWTATTWWTTTAAATTATWTTTTTTTTTAWTAWAWWTWAAAWA 522

Qy        867 GAAGTGGAAATAGATACCTTCACAAATCTTAAAAGACTTCTTCAGA 912
              : :  ::| | ||:  :|    : ||:| | : || || : |
Db        521 ATWAWWWWWAAAAATWWWWTWTTTTTWATTWATAAWTTTWTTTWWA 476
```

RESULT 6
AL565455/c
LOCUS AL565455 1201 bp mRNA linear EST 12-MAY-2003
DEFINITION AL565455 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone CS0DF005YO18 3-PRIME, mRNA sequence.
ACCESSION AL565455
VERSION AL565455.2 GI:30549492
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:12916848.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9232.f For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DF005BH09NP1&cluster=9232.f. Contact :

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 09:20:39 ; Search time 1031 Seconds
(without alignments)
11079.729 Million cell updates/sec

Title: US-09-430-412A-1
Perfect score: 3045
Sequence: 1 atcatcaataatatccgtta.....gtggaggatggggcttctcg 3045

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2475585 seqs, 1875730760 residues

Total number of hits satisfying chosen parameters: 4951170

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	428.2	14.1	559	15	US-10-027-632-196186	Sequence 196186,
c	2	428.2	14.1	559	15	US-10-027-632-196187	Sequence 196187,
	3	83.2	2.7	11745	14	US-10-240-453-206	Sequence 206, App
	4	82.4	2.7	6636	14	US-10-311-455-764	Sequence 764, App
	5	82	2.7	3673778	14	US-10-312-841-1	Sequence 1, Appli
c	6	80.8	2.7	158001	16	US-10-211-179-11	GENERAL INFORMATI
	7	78.8	2.6	19787	14	US-10-311-455-1424	Sequence 1424, Ap
	8	77.8	2.6	6063	14	US-10-240-453-268	Sequence 268, App
	9	75.2	2.5	6317	14	US-10-204-708-11	Sequence 11, Appl
	10	75.2	2.5	6317	14	US-10-311-455-381	Sequence 381, App
	11	75.2	2.5	6465	14	US-10-311-455-958	Sequence 958, App
	12	75.2	2.5	7351	14	US-10-311-455-2	Sequence 2, Appli
	13	74.8	2.5	5145	14	US-10-311-455-321	Sequence 321, App
	14	74.8	2.5	5145	14	US-10-240-485-17	Sequence 17, Appl
	15	74.6	2.4	3673778	14	US-10-312-841-2	Sequence 2, Appli
	16	74.4	2.4	14919	12	US-10-221-714A-228	Sequence 228, App
	17	74	2.4	5689	12	US-10-221-714A-148	Sequence 148, App
	18	74	2.4	5689	14	US-10-239-676-90	Sequence 90, Appl
	19	74	2.4	5689	14	US-10-240-453-100	Sequence 100, App
	20	74	2.4	12237	14	US-10-311-455-2331	Sequence 2331, Ap
	21	73.6	2.4	4985	14	US-10-094-240-10	Sequence 10, Appl
	22	73.6	2.4	4985	14	US-10-056-405-10	Sequence 10, Appl
	23	73.4	2.4	8170	14	US-10-240-453-132	Sequence 132, App
	24	73.2	2.4	18585	14	US-10-240-485-162	Sequence 162, App
c	25	72.6	2.4	863	15	US-10-027-632-31636	Sequence 31636, A
	26	72.6	2.4	960	14	US-10-198-846-6381	Sequence 6381, Ap
	27	72	2.4	5511	14	US-10-311-455-1844	Sequence 1844, Ap
	28	71.6	2.4	17594	14	US-10-311-455-2000	Sequence 2000, Ap
	29	71.4	2.3	8588	14	US-10-239-676-178	Sequence 178, App
	30	71.4	2.3	8588	14	US-10-240-453-200	Sequence 200, App
c	31	71.4	2.3	3673778	14	US-10-312-841-1	Sequence 1, Appli
c	32	71.2	2.3	5413	12	US-10-221-714A-418	Sequence 418, App
	33	71.2	2.3	19087	14	US-10-311-455-766	Sequence 766, App
c	34	71.2	2.3	276820	12	US-10-271-416-9	Sequence 9, Appli
	35	70.6	2.3	175590	10	US-09-911-077A-13	Sequence 13, Appl
	36	70.4	2.3	5388	14	US-10-311-455-218	Sequence 218, App
c	37	70.4	2.3	6247	14	US-10-311-455-248	Sequence 248, App
	38	70.2	2.3	7498	14	US-10-311-455-230	Sequence 230, App
	39	70.2	2.3	7823	14	US-10-239-676-198	Sequence 198, App
	40	70.2	2.3	7823	14	US-10-311-455-2034	Sequence 2034, Ap
	41	70.2	2.3	7823	14	US-10-240-453-292	Sequence 292, App
	42	69.8	2.3	3973	14	US-10-311-455-144	Sequence 144, App
	43	69.6	2.3	6192	12	US-10-221-714A-408	Sequence 408, App
	44	69.4	2.3	6298	14	US-10-239-676-64	Sequence 64, Appl
	45	69.4	2.3	6298	14	US-10-240-453-68	Sequence 68, Appl

ALIGNMENTS

RESULT 1
US-10-027-632-196186/c
; Sequence 196186, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 196186
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-196186

Query Match 14.1%; Score 428.2; DB 15; Length 559;
Best Local Similarity 92.2%; Pred. No. 2.7e-77;
Matches 517; Conservative 2; Mismatches 30; Indels 12; Gaps 6;

Qy 599 ATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTTGA 658

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 ATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTTGA 500

Qy        659 CACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTCCT 718
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 CACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTCCT 440

Qy        719 TTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTCAG 778
                |||||||||||  ||||||||||||||||||:||||||||||||||||||||||||||||||
Db        439 TTATTTTACTTGGTTTATAGTCTCAACTATAMTTTCAAAGTTTAATTTTAGATAATTCAG 380

Qy        779 CCTTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTTAACTGTAAAT 838
                 |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        379 -CTTTTAAATATTTTCCCATTATAATTTTTG-GACCTCTAACTCTATTTTAACTGTAAAT 322

Qy        839 ATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTT-A 897
                | ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        321 ACAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTTAA 262

Qy        898 AAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTGCAT 957
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db        261 AAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTTCAA 202

Qy        958 GCC----ATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTG-AATGATTGAGTGG 1012
                 |    |      || |  ||| |     ||||||||||||||| ||||||||||||
Db        201 TGCCATGAATCATCACAAATGGCAGGCTCATTGTGGTGGCATGCTGAAATGATTGAGTGG 142

Qy       1013 GACTGTGCC----AGCTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACACA 1068
                ||       |      ||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GAACTGTGCCCAGCACTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACACA 82

Qy       1069 CTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTGCT 1128
                ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db          81 CTGTCCTGTATTGTAATGCATTGGCCCAASTGGATTCTTTTTGATGCTTTGGTGATTGCT 22

Qy       1129 CTTTTGTTTGGGCTTGGAGAA 1149
                |||||||||||||||||||
Db          21 CTTTTGTTTGGGCTTGGAGAA 1
```

RESULT 2
US-10-027-632-196187/c
; Sequence 196187, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 196187
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-196187

Query Match 14.1%; Score 428.2; DB 15; Length 559;
Best Local Similarity 92.2%; Pred. No. 2.7e-77;
Matches 517; Conservative 2; Mismatches 30; Indels 12; Gaps 6;

```
Qy        599 ATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTTGA 658
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 ATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGTGTTGTTTGA 500

Qy        659 CACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTCCT 718
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 CACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTTAAGTTTCCT 440

Qy        719 TTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAGATAATTCAG 778
                |||||||||||  ||||||||||||||||||:||||||||||||||||||||||||||||||
Db        439 TTATTTTACTTGGTTTATAGTCTCAACTATAMTTTCAAAGTTTAATTTTAGATAATTCAG 380

Qy        779 CCTTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTTAACTGTAAAT 838
                 |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        379 -CTTTTAAATATTTTCCCATTATAATTTTTG-GACCTCTAACTCTATTTTAACTGTAAAT 322

Qy        839 ATAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTT-A 897
                |.|||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        321 ACAGTTCTGTATTTGTGAAGAGACTTTAGAAGTGGAAATAGATACCTTCACAAATCTTAA 262

Qy        898 AAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTGCAT 957
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db        261 AAAGACTTCTTCAGAGTCTGTAAACAGCATTACCATGTATACTTATCTCTTTCTTTTCAA 202

Qy        958 GCC----ATGATCATCACAATGCATGGCTCATGTGGTGGCATGCTG-AATGATTGAGTGG 1012
                 |    |      || |  ||| |     ||||||||||||||| ||||||||||||
Db        201 TGCCATGAATCATCACAAATGGCAGGCTCATTGTGGTGGCATGCTGAAATGATTGAGTGG 142

Qy       1013 GACTGTGCC----AGCTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACACA 1068
                ||       |      ||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GAACTGTGCCCAGCACTGAACTATAAAAAAAAAAAACAAACAAAACCTTATCCAAACACA 82

Qy       1069 CTGTCCTGTATTGTAATGCATTGGCCCAACTGGATTCTTTTTGATGCTTTGGTGATTGCT 1128
                ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db          81 CTGTCCTGTATTGTAATGCATTGGCCCAASTGGATTCTTTTTGATGCTTTGGTGATTGCT 22

Qy       1129 CTTTTGTTTGGGCTTGGAGAA 1149
                |||||||||||||||||||
Db          21 CTTTTGTTTGGGCTTGGAGAA 1
```

RESULT 3
US-10-240-453-206
; Sequence 206, Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associatec
; TITLE OF INVENTION: With DNA TRanscription
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240,453
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 206
; LENGTH: 11745

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 08:35:04 ; Search time 238 Seconds
(without alignments)
7100.107 Million cell updates/sec

Title: US-09-430-412A-1
Perfect score: 3045
Sequence: 1 atcatcaataatatccgtta.....gtggaggatggggcttctcg 3045

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	125.6	4.1	7218	1	US-08-232-463-14	Sequence 14, Appl
	2	75.2	2.5	6317	4	US-10-204-708-11	Sequence 11, Appl
c	3	74	2.4	665	2	US-08-883-795A-36	Sequence 36, Appl
	4	68.6	2.3	636	3	US-08-998-416-1137	Sequence 1137, Ap
	5	68.4	2.2	640681	4	US-09-790-988-1	Sequence 1, Appli
c	6	67.8	2.2	53526	3	US-08-658-136-2	Sequence 2, Appli
c	7	67.8	2.2	53577	3	US-08-658-136-1	Sequence 1, Appli
c	8	65.2	2.1	507	4	US-09-489-039A-1200	Sequence 1200, Ap
	9	65.2	2.1	516	4	US-09-489-039A-1064	Sequence 1064, Ap
	10	65.2	2.1	537	4	US-09-489-039A-1065	Sequence 1065, Ap
	11	65.2	2.1	549	4	US-09-489-039A-1130	Sequence 1130, Ap
c	12	65.2	2.1	558	4	US-09-489-039A-1011	Sequence 1011, Ap
c	13	65.2	2.1	609	4	US-09-489-039A-1177	Sequence 1177, Ap
	14	64.4	2.1	1866	3	US-09-173-581-13	Sequence 13, Appl
	15	64.4	2.1	1866	3	US-09-420-915-13	Sequence 13, Appl
c	16	63.8	2.1	19124	2	US-08-487-826B-13	Sequence 13, Appl
c	17	63.4	2.1	10640	4	US-09-417-485D-5	Sequence 5, Appli
	18	63.2	2.1	11049	4	US-10-204-708-21	Sequence 21, Appl
	19	62.8	2.1	819	4	US-09-918-686-5	Sequence 5, Appli
	20	62.8	2.1	92139	4	US-09-918-686-1	Sequence 1, Appli
	21	62.4	2.0	6866	4	US-10-204-708-20	Sequence 20, Appl
c	22	62.2	2.0	1268	4	US-09-369-247-42	Sequence 42, Appl
	23	62	2.0	615	3	US-08-998-416-186	Sequence 186, App
c	24	62	2.0	319608	4	US-09-539-333D-1	Sequence 1, Appli
c	25	62	2.0	319608	4	US-09-679-409-1	Sequence 1, Appli
c	26	61.8	2.0	1867	3	US-08-943-731-111	Sequence 111, App
c	27	61.8	2.0	24183	3	US-08-943-731-3	Sequence 3, Appli
c	28	61.8	2.0	202001	4	US-09-734-674-3	Sequence 3, Appli
	29	61.6	2.0	837	3	US-08-998-416-288	Sequence 288, App
	30	61.4	2.0	665	2	US-08-883-795A-36	Sequence 36, Appl
	31	60.6	2.0	685	4	US-09-183-266A-16	Sequence 16, Appl
	32	60.6	2.0	7122	3	US-09-318-448-2	Sequence 2, Appli
	33	60.6	2.0	7122	4	US-09-347-878-4	Sequence 4, Appli
	34	60.6	2.0	40000	4	US-09-780-049-18	Sequence 18, Appl
	35	60.2	2.0	5852	1	US-07-867-106-2	Sequence 2, Appli
	36	60.2	2.0	13205	4	US-09-835-811-3	Sequence 3, Appli
c	37	60.2	2.0	640681	4	US-09-790-988-1	Sequence 1, Appli
	38	59.8	2.0	8961	4	US-10-204-708-80	Sequence 80, Appl
c	39	59.6	2.0	20674	4	US-09-641-638-651	Sequence 651, App
	40	59.2	1.9	396	4	US-09-640-173-53	Sequence 53, Appl
	41	59.2	1.9	396	4	US-09-713-550-53	Sequence 53, Appl
	42	59	1.9	10467	4	US-10-204-708-2	Sequence 2, Appli
c	43	59	1.9	20303	1	US-08-370-975B-6	Sequence 6, Appli
c	44	59	1.9	26764	1	US-08-370-975B-1	Sequence 1, Appli
c	45	59	1.9	148567	4	US-09-801-876B-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 4.1%; Score 125.6; DB 1; Length 7218;

```
  Best Local Similarity   2.1%;  Pred. No. 4.3e-18;
  Matches    8;  Conservative  284;  Mismatches   88;  Indels    0;  Gaps    0;

Qy       2665 TTGGTCTGACCGGCAGGATCTGGTGTGTGTAAGTGAGTTCTGAGTCTCTGTTGACAAAAA 2724
Db       1441 TTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1382

Qy       2725 GAGACTCGAATGCAAAGACGCTGAGCTAGAGGGAGAGGAGGGCGGGGACCCAGAGGAAAG 2784
Db       1381 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1322

Qy       2785 AGGCACTCCTCGGGGTTGGGGAAGTATTAGGAGGGGAGGGTTAGAGTGGGAGGGAAGGAG 2844
Db       1321 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1262

Qy       2845 CCTGGCTTTCGAAGCGACTCACAGAGGGATAAATAAAGGGAAGTGAGGAGGAAGAGGGAG 2904
Db       1261 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1202

Qy       2905 ACTGAAAGGGAAGGCAGGTGGGGAGAAGGGGGACGAAAGAGGCAGAAGAGAGAGAAGAGA 2964
Db       1201 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1142

Qy       2965 GGAGGAGAGAGGGGGAGAGAGGGAAGGAAGGAAATAGGGAGAGGAGGGTCACAGAGTGAC 3024
Db       1141 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1082

Qy       3025 CGTGGAGGATGGGGCTTCTC 3044
Db       1081 RRRRRRRRRRRRRRRATCGC 1062
```

RESULT 2
US-10-204-708-11
; Sequence 11, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 11
; LENGTH: 6317
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-11

```
  Query Match            2.5%;  Score 75.2;  DB 4;  Length 6317;
  Best Local Similarity  45.0%;  Pred. No. 4.4e-07;
  Matches  391;  Conservative    5;  Mismatches  458;  Indels   14;  Gaps    3;

Qy          1 ATCATCAATAATATCCGTTATAAAGCTTGCTTTCTTTAGGTTAACTTTAGAGGCCTTGA 60
Db       1201 ATTGTAAGTATATTTTATTAATTTTTTTATTATGGTTTATATAGTATAGATAGTATTTTT 1260

Qy         61 AGAATAAGAGCTCATCTCTTTACAGGAGCTTTGGTTTGCAGCATTTACTTAAGAAATATT 120
Db       1261 ATTGGAGGATAATTTTTTATTAGAAAAGGATATTTAAATGAATTTTTTTAATAGAGGGAT 1320

Qy        121 TGGTATTCTGTATCTTTAAGAGTTAAACATAGAAGAATTGGCTAAGTGAAAATGAATGAA 180
Db       1321 TAATATTTGGTAGTAAGGAGTAAAAGAAAAAAAATTATATTATAAAAAAAAATTAAAATT 1380

Qy        181 ACGCAATATCATTCTGCATATATCATTTATTATATATCACAGTATTATTAGTTTTAAAAG 240
Db       1381 ATAAAATTAAGTTTATTTTATTTTAGTAAATATTGTTATAATTGTGATTATGTTATATTT 1440

Qy        241 TTAAACATAAATATCTATTATGYCATTGSACGAYTAGGYSAACCTA---RTCRGTGCTGC 297
Db       1441 TTTAGAAGAAATTTATTTTTATGTGTAATTTTTTTGATAAAATTATATGATAGTTATTT 1500

Qy        298 GAATACTTTCGATACTTCTGTTTCCCTCCTAGTATTCATAAGTGTGCCTTTGAAAACGTT 357
Db       1501 GTTAATATTTTTATTTTTGATTTTTGTAGAATAAATAATAAATGTTTTTGTAAAAGTTTT 1560

Qy        358 TTAAATTGTAAGAAATAAAATGTTTGATATATTATGTATATTATTACTAAGAAAAAACTT 417
Db       1561 TAGAATTTAAGTAATTTATTAAGTTTAAGTTTTTTTATTATTTTTTATGAAATATTTT 1620

Qy        418 GAATTACTTTGGATTTTGAAAAACTTTGATAAATTCTACATCATAGCATATTGAAGCAAG 477
Db       1621 AAATATATAGAGAAATTGAAAGAGTATTGTTAATAAATATATTAATGAATTAGATTTAAT 1680

Qy        478 AATAACAAATGCTATACCTCAGG--------AATATTAATTCCAGATTTTACAGCATTTT 529
Db       1681 AATTGTTAATATTTTGTCGTATGTATTTAATTATTTTAATTTTTTATGATTAATAATTT 1740

Qy        530 AACTTTCTTGATGAGAAAAAATAAATTTGTCAGTTATTAAACTATTTGGATCCAACAGAT 589
Db       1741 TAAAGTATTTATAAATATTATAAGATATTTTATTTTAAATATTTGATATGTATATTTT 1800

Qy        590 GAAAGCAGAATTCTAACTAACATATTTATTGATTTATTTGTGATTTACATATTTACATGT 649
Db       1801 AAGAGTAAGGATAAATTATATAGATATAATTATATTATAATATTTAAGAAAATTTTATAA 1860

Qy        650 GTTGTTTGACACAATTCTTAATTATGTTCTTGATATGCATATATTTGCTTCTTAAATTTT 709
Db       1861 TTTTTTAATATTATTTGAAAATTAATTTTTTTATTTTTATTTATTGTTTAATGTTTTTA 1920

Qy        710 AAGTTTCCTTTATTTTACTTTGTTTATAGTCTCAACTATAATTTCAAAGTTTAATTTTAG 769
Db       1921 TAGTTATTTTGTTTTGAATTAGATTTTAATTAAAGTTTATATATTGTATTTGGTTTTTAT 1980

Qy        770 ATAATTCAGCCTTTTAAATATTTTCCCATTATAATTTTTGTGACCTCTAACTCTATTTTA 829
Db       1981 ATAGT---GTTTTTGAGTTTTTTTTATTTTTTTTTTTTTTTTTTTATTATCGTTTTTATA 2037

Qy        830 ACTGTAAATATAGTTCTGTATTTGTGAA 857
Db       2038 GATTTATTGAGGTATAATTTATTGTAAA 2065
```

RESULT 3
US-08-883-795A-36/c
; Sequence 36, Application US/08883795A
; Patent No. 5985607
; GENERAL INFORMATION:
; APPLICANT: Delcuve, Genevieve
; APPLICANT: Awang, Gregor
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2